

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 20, 2006, 13:54:38 ; Search time 4201.73 Seconds
(without alignments)
16071.974 Million cell updates/sec

Title: US-10-634-548-18

Perfect score: 1188

Sequence: 1 gatcacaataatcttcaacac.....ccccaaaaaaaaaaaaa 1188

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenBank:*

1: gb_ba:*

SUMMARIES

Result No.	Score	Query length	DB ID	Description
1	1188	100.0	1188 6 E63071	E63071 DNA encodin
2	1188	100.0	1188 15 AF021244	AF021244 Arabidops
3	1175	98.9	1194 15 AY099815	AY099815 Arabidops
4	1128	94.9	1139 15 BT000309	BT000309 Arabidops
5	1115.4	93.9	1157 15 AY089091	AY089091 Arabidops
6	1012	85.2	90341 15 AC024609	AC024609 Arabidops
7	1012	85.2	119942 15 AC007797	AC007797 Arabidops
8	975	82.1	975 6 AX412268	AX412268 Sequence
9	975	82.1	975 6 AX412469	AX412469 Sequence
10	975	82.1	975 6 AX505463	AX505463 Sequence
11	754	63.5	1115 15 AF337544	AF337544 Brassica
12	395.8	33.3	453 10 BX510604	BX510604 Arabidops
13	239.4	20.2	987 6 AX411601	AX411601 Sequence
14	202.6	17.1	1174 6 AX411607	AX411607 Sequence
15	199	16.8	1242 6 AX411617	AX411617 Sequence
16	194.4	16.4	1104 6 BT009214	BT009214 Trifolium
17	188	15.8	1125 6 AX411609	AX411609 Sequence
18	188	15.8	1125 6 AX411611	AX411611 Sequence

19	178.8	15.1	1444 6 AX411613	AX411613 Sequence
20	176.8	14.9	1157 15 AY292526	AY292526 Ginkgo bi
21	169.8	14.3	1216 6 AX411603	AX411603 Sequence
22	169.4	14.3	1104 15 AF337545	AF337545 Brassica
23	169.2	14.2	992 15 AB056126	AB056126 Citrus un
24	167.6	14.1	1092 15 AF160869	AF160869 Citrus si
25	163.4	13.8	1135 6 E63072	E63072 DNA encodin
26	163.4	13.8	1135 15 AF134302	AF134302 Arabidops
27	161.8	13.6	1025 15 BT002898	BT002898 Arabidops
28	151.6	12.8	1340 6 E63070	E63070 DNA encodin
29	151.6	12.8	1340 15 AB025025	AB025025 Chenopodi
30	151.6	12.8	1340 15 AF134301	AF134301 Chenopodi
31	146.6	12.3	1159 15 BT013523	BT013523 Lycopersi
32	139.8	11.8	128855 14 AP007972	AP007972 Lotus cor
33	114.4	9.6	87286 15 AB026651	AB026651 Arabidops
34	110.8	9.3	1156 6 AX411625	AX411625 Sequence
35	110.8	9.2	1156 15 BT008923	BT008923 Trifolium
36	109.2	9.2	1302 6 AX411605	AX411605 Sequence
37	98.6	8.3	916 6 AX411615	AX411615 Sequence
38	98.6	8.3	916 15 BT009115	BT009115 Trifolium
39	82.6	7.0	1274 6 AX411627	AX411627 Sequence
40	73	6.1	884 15 AF337546	AF337546 Brassica
41	67	5.6	1300 6 E63080	E63080 DNA encodin
42	56.8	4.8	242 15 AT524543	AJ524543 Arabidops
43	56.6	4.8	110000 15 AP008208_107	Continuation (108
44	56.6	4.8	169378 15 AP004001	Continuation (108
45	53	4.5	110000 15 AP008216_142	Continuation (143

ALIGNMENTS

RESULT 1

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

ORIGIN

Query Match 100.0%; Score 1188; DB 6; Length 1188;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 GATCATAATATCTTCAACCACTTTTATTTATCTAGTTTATACAAATGGCGGATTA 60

Db	1	GATACATAAATCTTCAACAACTCTTTATATATCTAGTTTAAATACAAATGGCGCGATG	60
Qy	61	GAGGACAGTCCAAACGTTTTCTCTGTGTGTAACTCCGCGCGCTTTTGAGATAGACAGCTC	120
Db	61	GAGGACAGTCCAAACGTTTTCTCTGTGTGTAACTCCGCGCGCTTTTGAGATAGACAGCTC	120
Qy	121	CCGACAAACCGAGATACCGGTGGATCCGGTGGAAAATGATTCAACAGCACCGCAAAACCG	180
Db	121	CCGACAAACCGAGATACCGGTGGATCCGGTGGAAAATGATTCAACAGCACCGCAAAACCG	180
Qy	181	GTGAGATCACCTGTCCAAACAGTGGCGCGGAACTTATCCCGTGGTTTTATCTTCATGGC	240
Db	181	GTGAGATCACCTGTCCAAACAGTGGCGCGGAACTTATCCCGTGGTTTTATCTTCATGGC	240
Qy	241	TTTTATCTTCGCAACTACTCTTACTCTGACGTTCTTAAACAATGCGCTTCGACATGTTAC	300
Db	241	TTTTATCTTCGCAACTACTCTTACTCTGACGTTCTTAAACAATGCGCTTCGACATGTTAC	300
Qy	301	ATTCTTGTAGCCCAACAGTTGTCAAAATTTATCCGCGCGGAGGCGCACTGAAATGAGAC	360
Db	301	ATTCTTGTAGCCCAACAGTTGTCAAAATTTATCCGCGCGGAGGCGCACTGAAATGAGAC	360
Qy	361	GATGCTGGAAGTGAATTAACCTGGGCAATCGGAAACCTCAAGCTCACCTTACCACTTCG	420
Db	361	GATGCTGGAAGTGAATTAACCTGGGCAATCGGAAACCTCAAGCTCACCTTACCACTTCG	420
Qy	421	GTAATGTCTAATGAAAATATACACTCACCTGTGGGCAACAGCGCGGTGGAAAACGGCG	480
Db	421	GTAATGTCTAATGAAAATATACACTCACCTGTGGGCAACAGCGCGGTGGAAAACGGCG	480
Qy	481	TTTGGCGTTGGCGTACGCGCAATCGCGCAATTAAGACCATCATCACTTCACTCTTA	540
Db	481	TTTGGCGTTGGCGTACGCGCAATCGCGCAATTAAGACCATCATCACTTCACTCTTA	540
Qy	541	ATAGGAATGATCCAGTCGCGAGGAATTAACAATTAATAGAACCGATCCGATATCTTA	600
Db	541	ATAGGAATGATCCAGTCGCGAGGAATTAACAATTAATAGAACCGATCCGATATCTTA	600
Qy	601	ACGTATAAACCGGAATCTTTGAGCTGGAACATACCGTTGCACTGTGGGAAACCGGACTC	660
Db	601	ACGTATAAACCGGAATCTTTGAGCTGGAACATACCGTTGCACTGTGGGAAACCGGACTC	660
Qy	661	GGAACCGAAGTGGAAACAACGTATGCAATGCGCAACCAACGGAATTAACCATAGAGAG	720
Db	661	GGAACCGAAGTGGAAACAACGTATGCAATGCGCAACCAACGGAATTAACCATAGAGAG	720
Qy	721	TTTTTCAAAAGAGTCTAAGCGCAACGAACCCATTTTCGTGGCTGCGGATTAACGATATG	780
Db	721	TTTTTCAAAAGAGTCTAAGCGCAACGAACCCATTTTCGTGGCTGCGGATTAACGATATG	780
Qy	781	GATATGTTGACAGATATTTGCGCGGTTTTGTTGGGTTTATGCGCGGTTGATGTGTAG	840
Db	781	GATATGTTGACAGATATTTGCGCGGTTTTGTTGGGTTTATGCGCGGTTGATGTGTAG	840
Qy	841	AATGGCAAGAAAAAGCTGAGATGAGAGCTTTGTAGGTGAATGTGTGGTCGTT	900
Db	841	AATGGCAAGAAAAAGCTGAGATGAGAGCTTTGTAGGTGAATGTGTGGTCGTT	900
Qy	901	CTCAAGTATAGTTGTGGGGTGAATAAGCGAGATTCCATGATTTGTGAAGATCCTTCC	960
Db	901	CTCAAGTATAGTTGTGGGGTGAATAAGCGAGATTCCATGATTTGTGAAGATCCTTCC	960
Qy	961	GTTTCTCCGGCCAACTTGATCCTTCACTGAGTTGAAAGAAAGCTTCGTGGTATCTTCGTC	1020
Db	961	GTTTCTCCGGCCAACTTGATCCTTCACTGAGTTGAAAGAAAGCTTCGTGGTATCTTCGTC	1020
Qy	1021	TAGATTTGTATGTACTATATATCAGAGGGGCTTGAATATTTGAAAAAACCATCAATG	1080
Db	1021	TAGATTTGTATGTACTATATATCAGAGGGGCTTGAATATTTGAAAAAACCATCAATG	1080
Qy	1081	TTTTCTAGCTCCAACTGACTATTTGTTCAATGCTTAAGTGCATGTGATTTTTTATTA	1140
Db	1081	TTTTCTAGCTCCAACTGACTATTTGTTCAATGCTTAAGTGCATGTGATTTTTTATTA	1140

DB	1081	TTTTCAGCTCCGAAGCTAGCACTATTGTCAGTCTCTTAAGTGCATGTATTTTATTTAA	1140
QY	1141	CTGCATCAAAACATTGTGTATAGTTTACCACCAAAAAAAAAAAAAAAAA	1188
DB	1141	CTGCATCAAAACATTGTGTATAGTTTACCACCAAAAAAAAAAAAAAAAA	1188
RESULT 2			
AF021244			
LOCUS			
DEFINITION	Arabidopsis thaliana coronatine-induced protein 1 (COR11) mRNA,		
ACCESSION	AF021244		
VERSION	AF021244.1		
KEYWORDS	GI:2460202		
SOURCE	Arabidopsis thaliana (thale cress)		
ORGANISM	Arabidopsis thaliana		
REFERENCE	1. (bases 1 to 1188)		
AUTHORS	Benedicti,C.E., Costa,C.L., Turcinelli,S.R. and Arruda,P.		
TITLE	Differential expression of a novel gene in response to coronatine, methyl jasmonate, and wounding in the Col mutant of Arabidopsis		
JOURNAL	Plant Physiol. 116 (3), 1037-1042 (1998)		
PIBMED	9501136		
REFERENCE	2. (bases 1 to 1188)		
AUTHORS	Benedicti,C.E., Costa,C.L., Turcinelli,S.R. and Arruda,P.		
TITLE	Direct Submission		
JOURNAL	Submitted (26-AUG-1997) Centro de Biologia Molecular e Engenharia Genetica, University of Campinas, Campinas, SP 13083970, Brazil		
FEATURES	Location/Qualifiers		
source	1. .1188		
gene	/organism="Arabidopsis thaliana"		
	/mol_type="mRNA"		
	/cultivar="Columbia; Landsberg erecta"		
	/db_xref="taxon:3702"		
	/tissue_type="seedlings; leaves; flowers; siliques"		
	1. .1188		
	/gene="COR11"		
	/note="ATHCOR1; coronatine-induced gene; induced in response to the phytoalexin coronatine, methyl jasmonate and wounding"		
	49. .1023		
	/gene="COR11"		
	/note="similarity to certain plant and animal hydrolases; includes a P-loop, a potential ATP/GMP binding site"		
	/codon_start=1		
	/product="coronarine-induced protein 1"		
	/protein_id="AAC13947.1"		
	/db_xref="GI:2460203"		
	/translation="MAAIEDSPFSSVVTAAFEISGLPTTEIPVDVENDSTAPRKRVRITPCVAGTAVPVLPFFHGPVLRNYSVIVLHSHGIVLVAPOCLKLPQGOVEVDDAGSTNNASERLKAHLPTSVNRANKSTSLYHSGSGKTAPRALVAGHAATLDPSTTFSLALGIDVPAGTNRK1TRIDPHILITYKPSFELDIPLVAVGTGSPKNNVMPCAPTFDNLHEEYKCKAKATKAFVADYGHMDMDDDLPGLGFVAGGCKCKNGQRKSEMSFVGGIVAVFLKYSIMGEKAEIRLLIVKDPSPVPAKLDPSPPEBESGIPV"		
ORIGIN			
Query Match	100.0%;	Score 1188;	DB 15; Length 1188;
Best Local Similarity	100.0%;	Pred. No. 0;	
Matches 1188;	Conservative 0;	Mismatches 0;	Indels 0; Gaps 0;
QY	1	GATACATTAATCTTCACACAACTCTTAATTATCTAGTTTAATACAAATGGCGCGTA	60
DB	1	GATACATTAATCTTCACACAACTCTTAATTATCTAGTTTAATACAAATGGCGCGTA	60
QY	61	GAGGACAGTCCAAAGCTTTCTCTGTGTACTCCGCGGCTTTTGATAGGAGCTC	120
DB	61	GAGGACAGTCCAAAGCTTTCTCTGTGTACTCCGCGGCTTTTGATAGGAGCTC	120
QY	121	CCGACAAACGAGATACCGGTGGATCCGGTGGAAATGATTCACAGACCGCCAAACG	180

Db 121 CCGAACCGAGATACCGGTGATCCGGTGGAAAAATGATTCAGACGACCGCCAAAACCG 180
 QY 181 GTGAGATCACTGTCCCAAGTGGCGGAACTTATCCGCTGTTTATTTCTTCATGCG 240
 Db 181 GTGAGATCACTGTCCCAAGTGGCGGAACTTATCCGCTGTTTATTTCTTCATGCG 240
 QY 241 TTTTATCTGGCACTCTTACTCTGAGCTTCTTAACCAATCGCTGGCATGCTTAC 300
 Db 241 TTTTATCTGGCACTCTTACTCTGAGCTTCTTAACCAATCGCTGGCATGCTTAC 300
 QY 301 ATTCTTGTAGCCCAAGTGTGTGAATTTATGCGCGCGGAGGCAAGTGAAGTGAAC 360
 Db 301 ATTCTTGTAGCCCAAGTGTGTGAATTTATGCGCGCGGAGGCAAGTGAAGTGAAC 360
 QY 361 GATCTGGAAGTGTGAATTAATCTGGGCAATCGGAAAACCTCAAGCTCACTCAACTTCG 420
 Db 361 GATCTGGAAGTGTGAATTAATCTGGGCAATCGGAAAACCTCAAGCTCACTCAACTTCG 420
 QY 421 GTAAATGTAAATGAAAAATACACTCACTGCGGGGCAAGCGCGGGTGGAAAAACGGCG 480
 Db 421 GTAAATGTAAATGAAAAATACACTCACTGCGGGGCAAGCGCGGGTGGAAAAACGGCG 480
 QY 481 TTTCGGGTGGCGTACGCGCATCGCGCAATTAAGCCATCATCAGTTTTCAGCTCTA 540
 Db 481 TTTCGGGTGGCGTACGCGCATCGCGCAATTAAGCCATCATCAGTTTTCAGCTCTA 540
 QY 541 ATAGGAATGTATCCAGTGGCAAGAACTTAACAATTAATTAAGAACCGATCCGATATCTTA 600
 Db 541 ATAGGAATGTATCCAGTGGCAAGAACTTAACAATTAATTAAGAACCGATCCGATATCTTA 600
 QY 601 ACGATTAACCGGAATCTTTCAGACTGGACATACCGGTGAGTGGAGGAAACGGGACTC 660
 Db 601 ACGATTAACCGGAATCTTTCAGACTGGACATACCGGTGAGTGGAGGAAACGGGACTC 660
 QY 661 GGAACCGAAGTGAACAAGTGTATGCGCAATGCGCAACAAGCACTTAACCAATGAGAG 720
 Db 661 GGAACCGAAGTGAACAAGTGTATGCGCAATGCGCAACAAGCACTTAACCAATGAGAG 720
 QY 721 TTTTACCAAGAGTGAAGGCGAAGAAACCATTTCTGCTGCGGATTAACGACATATG 780
 Db 721 TTTTACCAAGAGTGAAGGCGAAGAAACCATTTCTGCTGCGGATTAACGACATATG 780
 QY 781 GATATGTGAGCATGATTTGCGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 840
 Db 781 GATATGTGAGCATGATTTGCGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 840
 QY 841 AATGGCAAAAGAAAAAGTGTGAGATGAGAGCTTTGTAGGTGGAATTTGTGCGCTT 900
 Db 841 AATGGCAAAAGAAAAAGTGTGAGATGAGAGCTTTGTAGGTGGAATTTGTGCGCTT 900
 QY 901 CTCAAGTATAGTTTGTGGGGTGAAGAAACCGGAGATTCGATTTGTAAGAGATCTTCC 960
 Db 901 CTCAAGTATAGTTTGTGGGGTGAAGAAACCGGAGATTCGATTTGTAAGAGATCTTCC 960
 QY 961 GTTTCTCCGGCAAGCTTGAATCTTCACTGAGTGAAGAAAGCTTCTGATCTTCTGTC 1020
 Db 961 GTTTCTCCGGCAAGCTTGAATCTTCACTGAGTGAAGAAAGCTTCTGATCTTCTGTC 1020
 QY 1021 TAGATTTGTATGTATGATATTAATCAAGAGGCTTGAATTTGTAAGAAACCTATCATG 1080
 Db 1021 TAGATTTGTATGTATGATATTAATCAAGAGGCTTGAATTTGTAAGAAACCTATCATG 1080
 QY 1081 TTTTCTAGCTTCAAGCTATGATTTGCTCAATGCTTAAGTTCATGTTTATTTATTA 1140
 Db 1081 TTTTCTAGCTTCAAGCTATGATTTGCTCAATGCTTAAGTTCATGTTTATTTATTA 1140
 QY 1141 CTGATCAAAACATTTGTATAGTTTAAACCCCAAAAAA 1188
 Db 1141 CTGATCAAAACATTTGTATAGTTTAAACCCCAAAAAA 1188

LOCUS AY099815 1194 bp mRNA linear PLN 06-MAY-2002
 DEFINITION Arabidopsis thaliana unknown protein (At1g19670) mRNA, complete cds.
 ACCESSION AY099815
 VERSION AY099815.1 GI:20466697
 KEYWORDS FLI CDNA.
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 1 (bases 1 to 1194)
 REFERENCE Nguyen,M., Karlin-Neumann,G., Southwick,A., Lam,B., Miranda,M., Palm,C.J., Bower,L., Jones,T., Banu,O., Carninci,P., Chen,H., Cheuk,R., Chung,M.K., Hayashizaki,Y., Ishida,J., Kamiya,A., Kawai,J., Kim,C., Lin,J., Liu,S.X., Narusaka,M., Pham,P.K., Sakano,H., Sakurai,T., Satou,M., Seki,M., Shimizu,P., Yamada,K., Shinzaki,K., Becker,J., Theologis,A. and Davis,R.W.
 Direct Submission
 Submitted (24-APR-2002) DNA Sequencing and Technology Center, Stanford University, 855 California Avenue, Palo Alto, CA 94304, USA
 TITLE JOURNAL
 COMMENT e-mail for correspondence: arabseq.stanford.edu
 RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA : "RIKEN Arabidopsis Full-length cDNA") : Seki,M., Narusaka,M., Ishida,J., Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J., Hayashizaki,Y. and Shinzaki,K.
 The Salk, Stanford, REEC (SSP) Consortium members carried out the sequencing and annotation of the RAFL cDNAs : Nguyen,M., Southwick,A., Karlin-Neumann,G., Lam,B., Miranda,M., Palm,C.J., Bower,L., Jones,T., Banu,O., Chen,H., Cheuk,R., Chung,M.K., Kim,C., Lin,J., Liu,S.X., Pham,P.K., Sakano,H., Shimizu,P., Yamada,K., Becker,J., Theologis,A. and Davis,R.W.
 Nguyen,M. (SSP/Stanford) and Seki,M. (RIKEN GSC) contributed equally to this work. Shinzaki,K. (RIKEN GSC) and Davis,R.W. (SSP/Stanford) contributed equally to this work as PIs.
 FEATURES
 source location/Qualifiers
 1..1194
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /db_xref="taxon:3702"
 /chromosome="1"
 /clone="RAFL09-95-F15 (R25648)"
 /ecotype="Columbia"
 /note="This clone is in pBluescript"
 1..1194
 /gene="At1g19670"
 /note="synonym: F14P1.11"
 48..1022
 /gene="At1g19670"
 /codon_start=1
 /product="unknown protein"
 /protein_id="AA020666.1"
 /db_xref="GI:20466698"
 /translation="MAAIDSPSPSSVTPPAFEISLPTEIPVDPVENDSTAPPKPVRITCPVAGTVPVLFPHGYLRNRYFSDVNLHLSHYIIAVALGCLLPKGGVGVNDAGSVIMWASBNLKAHLPTSVNMGKXTSLVGRSGGKTAAYVALGHAATLDPSSITFSLGIDIPVAGTNRKIRIDPHILTYKSPSEFLDIPVAVYVGLGKRNKNNVPPCAPTDLHNEFYECRATKAHFAVADYGHMDLDDDDIPGVGMACGMCNCKRKSSEMSFVGDIIVAFKYSIMGKABIRLIVDPSVSPAKLPDPSPELBSGIFV"
 ORIGIN
 Query Match 98.9%; Score 1175; DB 15; Length 1194;
 Best Local Similarity 100.0%; Pred.No.0;
 Matches 1175; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 2 ATACATTAATCTTCAACACACTCTTAATTAATCTAGTTAATATACAAATGGCGCATAG 61
 1 ATACATTAATCTTCAACACACTCTTAATTAATCTAGTTAATATACAAATGGCGCATAG 60

QY 62 AGAAGAGTCCAAAGTTTCTCTGTGTGTAACCTCCGGCGGCTTTTGAGATAGGACGCTCC 121
 DB 61 AGGAGAGTCCAAAGTTTCTCTGTGTGTAACCTCCGGCGGCTTTTGAGATAGGACGCTCC 120
 QY 122 CGACAAACCGAGATACCGGTGATCCGGTGAAATGATTCACAGCACCGCCAAACCGG 181
 DB 121 CGACAAACCGAGATACCGGTGATCCGGTGAAATGATTCACAGCACCGCCAAACCGG 180
 QY 162 TGAGAAATACCTGTCCAAACAGTCGCGGAACTTAATCCGTGTTTATTTCCATGGCT 241
 DB 161 TGAGAAATACCTGTCCAAACAGTCGCGGAACTTAATCCGTGTTTATTTCCATGGCT 240
 QY 242 TTTATCTTGCAGTACTTCTACTCTGACGTTCTTAACCAACATCGCTTCGATGTTTCA 301
 DB 241 TTTATCTTGCAGTACTTCTACTCTGACGTTCTTAACCAACATCGCTTCGATGTTTCA 300
 QY 302 TTTCTGTAGCCCAAGTTGTGCAAAATTAATGCGCGCGGAGGCAAGTGAAGTGAAG 361
 DB 301 TTTCTGTAGCCCAAGTTGTGCAAAATTAATGCGCGCGGAGGCAAGTGAAGTGAAG 360
 QY 362 ATGCTGGAAGTGTGTAATCTGGGATCGGAAACCTCAAGCTCACTTACCAACTTGG 421
 DB 361 ATGCTGGAAGTGTGTAATCTGGGATCGGAAACCTCAAGCTCACTTACCAACTTGG 420
 QY 422 TAAATGCTAATGAAATAACACCTCACTGTGGGCAAGCGCGGTGGGAAACGGCGT 481
 DB 421 TAAATGCTAATGAAATAACACCTCACTGTGGGCAAGCGCGGTGGGAAACGGCGT 480
 QY 482 TTGCGGTGCGTAGGCGCAATGCGCAACATTAGACCATCATCATCGTTTCACTTAA 541
 DB 481 TTGCGGTGCGTAGGCGCAATGCGCAACATTAGACCATCATCATCGTTTCACTTAA 540
 QY 542 TAGGAATTGATCCATCGCAGGAACATAAATACATTAGAACCGATCCGATATCTTAA 601
 DB 541 TAGGAATTGATCCATCGCAGGAACATAAATACATTAGAACCGATCCGATATCTTAA 600
 QY 602 CGTATTAACCGGAACTTTTCAGAGTGAACATACCGGTTGAGTGTGGAAACCGGACTCG 661
 DB 601 CGTATTAACCGGAACTTTTCAGAGTGAACATACCGGTTGAGTGTGGAAACCGGACTCG 660
 QY 662 GACCGAAGTGAACCAAGTATGCCAATGCGGACCAACGAGCTTAAACCATGAGAGT 721
 DB 661 GACCGAAGTGAACCAAGTATGCCAATGCGGACCAACGAGCTTAAACCATGAGAGT 720
 QY 722 TTTAAGAGTGTAGGCGAAGGAAAGCCATTTGTGTGCGGATTAACGACATATAG 781
 DB 721 TTTAAGAGTGTAGGCGAAGGAAAGCCATTTGTGTGCGGATTAACGACATATAG 780
 QY 782 ATATGTGAGACGATGATTTGCGCGGTTTGTGTGTTATGCGCGGTTGATGTTAGA 841
 DB 781 ATATGTGAGACGATGATTTGCGCGGTTTGTGTGTTATGCGCGGTTGATGTTAGA 840
 QY 842 ATGGCAAAAGAAAAGTCTGAGATGAGAGCTTTGTAGTGAATGTGTTGCGTTTC 901
 DB 841 ATGGCAAAAGAAAAGTCTGAGATGAGAGCTTTGTAGTGAATGTGTTGCGTTTC 900
 QY 902 TCAAGTATAGTTTGTGGGGTGAAGGAGGAGATTGATGATGTGAAGATCCTTCG 961
 DB 901 TCAAGTATAGTTTGTGGGGTGAAGGAGGAGATTGATGATGTGAAGATCCTTCG 960
 QY 962 TTTCTCCGCGCAAGCTTGAATCTTCACTGAGTTGGAAGAGCTTCTGTATCTTCGCT 1021
 DB 961 TTTCTCCGCGCAAGCTTGAATCTTCACTGAGTTGGAAGAGCTTCTGTATCTTCGCT 1020
 QY 1022 AGATTGTGTTATGTATATATACAGAGGGGCTTGAATATTTGAAAAACCTATCATGT 1081
 DB 1021 AGATTGTGTTATGTATATATACAGAGGGGCTTGAATATTTGAAAAACCTATCATGT 1080
 QY 1082 TTTTCAAGTCCAAAGTACGATTTGTCAATGTCCTAAGTGTGATGTTTATATTAAC 1141
 DB 1081 TTTTCAAGTCCAAAGTACGATTTGTCAATGTCCTAAGTGTGATGTTTATATTAAC 1140

QY 1142 TCGATCAAAACATTTGTATGTTTACCCCAAAA 1176
 DB 1141 TCGATCAAAACATTTGTATGTTTACCCCAAAA 1175
 RESULT 4
 BT000309
 LOCUS
 DEFINITION
 cde.
 BT000309 1139 bp mRNA linear PLN 19-SEP-2002
 Arabidopsis thaliana unknown protein (Atlg19670) mRNA, complete
 ACCESSION
 BT000309.1 GI:23198201
 VERSION
 BT000309.1
 KEYWORDS
 SOURCE
 ORGANISM
 Arabidopsis thaliana (thale cress)
 Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 1 (bases 1 to 1139)
 REFERENCE
 Nguyen,M., Karlin-Neumann,G., Southwick,A., Tripp,M., Miranda,M.,
 Palm,C.J., Bowser,L., Jones,T., Banb,J., Carninci,P., Chen,H.,
 Cheuk,R., Chung,M.K., Hayashizaki,Y., Ishida,J., Kamiya,A.,
 Kawai,J., Kim,C., Lin,J., Liu,S.X., Natusaka,M., Pham,P.K.,
 Sakano,H., Sakurai,T., Saitou,M., Seki,M., Shimu,P., Yamada,K.,
 Shinozaki,K., Becker,J., Theologis,A. and Davis,R.W.
 TITLE
 Submitted (19-SEP-2002) DNA Sequencing and Technology Center,
 JOURNAL
 Stanford University, 855 California Avenue, Palo Alto, CA 94304,
 USA
 COMMENT
 e-mail for correspondence: arab@sequence.stanford.edu
 The Salk, Stanford, PGSC (SSP) Consortium members constructed and
 sequenced the PENTR (ORF) clones using the RAPL cDNAs: Nguyen,M.,
 Southwick,A., Tripp,M., Palm,C.J., Jones,T., Wu,T., Chen,H.,
 Cheuk,R., Chan,M.M., Chang,C.H., Dale,J.M., Deng,J.M., Hsuan,Y.W.,
 Lee,J.M., Kim,C.J., Quach,H.L., Shinn,P., Tang,C.C., Torouni,M.,
 Wallender,E.K., Wong,C., Wu,H.C., Yamada,K., Yu,G., Yuan,S.,
 Becker,J., Theologis,A. and Davis,R.W.
 Nguyen,M., (SSP/Stanford) and Seki,M. (RIKEN GSC) contributed
 equally to this work. Shinozaki,K. (RIKEN GSC) and Davis,R.W.
 (SSP/Stanford) contributed equally to this work as PIs.
 FEATURES
 source
 1. 1139
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /db_xref="taxon:3702"
 /chromosome="1"
 /clone="U25648"
 /ecotype="Columbia"
 /note="This clone is in PENTR/SD-dTPO. This is a cloned
 PCR product using RIKEN clone RAPL09-95-F15 (AT099815) as
 a template"
 1. 1139
 /gene="Atlg19670"
 /note="synonym: F14P1.11"
 1. 975
 /gene="Atlg19670"
 /codon_start=1
 /product="unknown protein"
 /protein_id="AA115628.1"
 /db_xref="GI:23198202"
 /translation="MAIEDSPTESSVTPPAFEIGSLPTTEIPVDPEVENDSTAPPK
 VRTCPVAGTVPVLPFHGFLRVNFYSVNLVHSHGYILVAPOLCKLIPGCGOVE
 VDDGQSYTNASENLKHLPTSYNANGKTSLYGSHSGGKTARVALGHAATDPSIT
 FSLALIGDIPVAGTNKYLITDPLILTYKPSSEFLDIPVAVGTGIGPWNVMPRCAPT
 DLNHEEFYKCKATKAHPVADYGHMDLDDLPFVFGMAGCMCKXGQRKSEMRSP

ORIGIN VGGIIVAFLKXSLWGEKAEIRLIIVKDPSPVSPAKDPSPELEASGIFV"

Query Match 94.9%; Score 1128; DB 15; Length 1139;
Best Local Similarity 100.0%; Pred. No. 1.6e-306;
Matches 1128; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

49 ATGGCGCGGATGAGGACAGTCCAAAGCTTTCTCTGTGTGTAAGTCCGCGGCTTTTGGAG 108
1 ATGGCGCGGATGAGGACAGTCCAAAGCTTTCTCTGTGTGTAAGTCCGCGGCTTTTGGAG 60
109 ATAGGAGAGCTCCCGGACAAACCGAGATACCGGTGATCCGGTGGAAAAATGATTCAACAGCA 168
61 ATAGGAGAGCTCCCGGACAAACCGAGATACCGGTGATCCGGTGGAAAAATGATTCAACAGCA 120
169 CCGCCAAACCGGTGAGAAATCAGTTCGCAACAGTCCCGGAACTTATCCCGTGTGTTTA 228
121 CCGCCAAACCGGTGAGAAATCAGTTCGCAACAGTCCCGGAACTTATCCCGTGTGTTTA 180
229 TTCTTCATGAGCTTTTATCTTGGCACTACTCTTACTCTGACGTTCTTTAACCAATGCTT 288
181 TTCTTCATGAGCTTTTATCTTGGCACTACTCTTACTCTGACGTTCTTTAACCAATGCTT 240
289 TCGCATGATGATCTTTCTTGTAGCCCAAGTTGTGCAAAATATTGCGCGGAGGAGCA 348
241 TCGCATGATGATCTTTCTTGTAGCCCAAGTTGTGCAAAATATTGCGCGGAGGAGCA 300
349 GTGGAAGTGAAGATGCTGGAAGTGTGATTAATCTGGGATCGAAAACTTCAAAAGCTCAC 408
301 GTGGAAGTGAAGATGCTGGAAGTGTGATTAATCTGGGATCGAAAACTTCAAAAGCTCAC 360
409 CTACCAACTTCGGTAATTCCTTAATGAAAAATACCTCAGTGTGGGCGCACAGCCGGGT 468
361 CTACCAACTTCGGTAATTCCTTAATGAAAAATACCTCAGTGTGGGCGCACAGCCGGGT 420
469 GGGAAAAAGCGGCTTGGCGGTGCGCTAGAGCCATCCCGCAACATTAGACCCATCAGC 528
421 GGGAAAAAGCGGCTTGGCGGTGCGCTAGAGCCATCCCGCAACATTAGACCCATCAGC 480
529 TTTTCAGCTCTAATAGAAATGATCCAGTGCAGAACTTAACAAATACATTAGAACCGAT 588
481 TTTTCAGCTCTAATAGAAATGATCCAGTGCAGAACTTAACAAATACATTAGAACCGAT 540
589 CCGCATATCTTAAAGTATTAACCGGAATCTTTGAGCTGACATACCGGTTGCAATGCTG 648
541 CCGCATATCTTAAAGTATTAACCGGAATCTTTGAGCTGACATACCGGTTGCAATGCTG 600
649 GGAACCGGACTCGGACCGAAGTGAACAGTGTGCAACCATGCGACCAACGGAAGCTTA 708
601 GGAACCGGACTCGGACCGAAGTGAACAGTGTGCAACCATGCGACCAACGGAAGCTTA 660
709 AACCATGAGAGATTTTTACAAAGAGTGTAGGCGACGAAAGCCATTTCTGTGCTGCGGAT 768
661 AACCATGAGAGATTTTTACAAAGAGTGTAGGCGCGAAGGCAATTTCTGTGCTGCGGAT 720
769 TACGGAATATGATATGATGAGAGATTTTGGCGGTTTGTGTGGGTTTATGCGCGGT 828
721 TACGGAATATGATATGATGAGAGATTTTGGCGGTTTGTGTGGGTTTATGCGCGGT 780
829 TGTATGTGTAAGATGAGGCAAAAGAAAGTCTGAGATGAGAGCTTTGTAGGTGAAT 888
781 TGTATGTGTAAGATGAGGCAAAAGAAAGTCTGAGATGAGAGCTTTGTAGGTGAAT 840
889 GTGTGTCGTTTCTCAAGATATAGTTGTGTGGGTGAAAAAGCGAGATTCAGATTGATG 948
841 GTGTGTCGTTTCTCAAGATATAGTTGTGTGGGTGAAAAAGCGAGATTCAGATTGATG 900
949 AAGGATCTTCCGTTTCTCCGCGCAAGCTTATCTTCACTGATGTTGGAAGAGCTTCT 1008
901 AAGGATCTTCCGTTTCTCCGCGCAAGCTTATCTTCACTGATGTTGGAAGAGCTTCT 960
1009 GGTATCTTCTGATGATTTGTGATATGATATATCAGAGGGGTCTTGAATATTTGAAA 1068

Db 961 GGTATCTTCTGATGATTTGTGATGATATATCAGAGGGGTCTTGAATATTTGAAA 1020
QY 1069 AACCTATCAATGTTTCTTACTCCCAAGCTAGTATGTTGATGCTTAAGTCAATGTGT 1128
Db 1021 AACCTATCAATGTTTCTTACTCCCAAGCTAGTATGTTGATGCTTAAGTCAATGTGT 1080
QY 1129 ATTTTATTAATCACTGATCAAAACATTTGTTATAGTTTATCCCAAAA 1176
Db 1081 ATTTTATTAATCACTGATCAAAACATTTGTTATAGTTTATCCCAAAA 1128

RESULT 5
AY089091 1157 bp mRNA linear PLN 14-APR-2003
LOCUS
DEFINITION Arabidopsis thaliana clone 31589 mRNA, complete sequence.
ACCESSION
VERSION AY089091.1 GI:21407865
KEYWORDS
SOURCE
ORGANISM Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.

REFERENCE
AUTHORS Haas,B.J., Volfovsky,N., Town,C.D., Troupkan,M., Alexandrov,N.,
Feldmann,K.A., Flavell,R.B., White,O. and Salzberg,S.L.
Full-length messenger RNA sequences greatly improve genome
annotation
Genome Biol. 3 (6), RESEARCH0029 (2002)

TITLE
JOURNAL
PUBMED 12093376
REFERENCE
AUTHORS 2 (bases 1 to 1157)
Feldmann,K., Troupkan,M., Alexandrov,N., Lu,Y.-P., Flavell,R. and
Feldmann,K.
Full-length cDNA from Arabidopsis thaliana
Unpublished
3 (bases 1 to 1157)
Brover,V., Troupkan,M., Alexandrov,N., Lu,Y.-P., Flavell,R. and
Feldmann,K.

TITLE
JOURNAL
COMMENT Direct Submission
Submitted (11-MAR-2002) Ceres, Inc, 3007 Malibu Canyon Road,
Malibu, CA 90265, USA
This clone sequence is one of 5,000 Ceres full-length cDNAs made
available to TIGR and Genbank. The following quality assessment of
this set was done by comparison with known proteins: two percent of
the clones are estimated to be 5'-truncated; less than one percent
are 3'-truncated; approximately two percent represent alternative
splice variants, including unspliced introns and spliced exons; one
percent may contain premature stop codons; five percent may have
frame shifts in a coding region. A sequence is considered to be
5'-truncated if it lacks the translation initiation start (ATG). A
sequence is considered to be 3'-truncated if it lacks the
C-terminal end of the encoded protein. Please note that these cDNA
sequences are derived from the Ws or Laer ecotypes and therefore
may contain polymorphisms when compared to sequences from Col-0.
Genest carried out the library production and sequencing of the
full-length clones. Ceres, Inc, carried out the clustering of the
5' sequences, selection of clones, and sequence assembly.

FEATURES
source
1..1157
Location/Qualifiers
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="31589"

ORIGIN

Query Match 93.9%; Score 115.4; DB 15; Length 1157;
Best Local Similarity 99.5%; Pred. No. 5.8e-303;
Matches 1151; Conservative 0; Mismatches 1; Indels 5; Gaps 3;
QY 2 ATACATTAATCTTCAACACAACTTTAATTAATGATTAAATACAAATGGCGGAGTAG 61
Db 5 ATACATTAATCTTCAACACAACTTTAATTAATGATTAAATGAAATGGCGGAGTAG 64

QY 62 AGGACAGTCCAACTTTTCTCTGTGTAACTCCGGCGGCTTTTGAGATAGGACGCTCC 121
 DB 65 AGGACAGTCCAACTTTTCTCTGTGTAACTCCGGCGGCTTTTGAGATAGGACGCTCC 124
 QY 122 CGACACCGAGATACCGGTGATCCGGTGGAAATGATTCAACAGACCGCCAAACCGG 181
 DB 125 CGACACCGAGATACCGGTGATCCGGTGGAAATGATTCAACAGACCGCCAAACCGG 184
 QY 182 TGAGATACCTGTCCAAAGTCCGGGAACTTATCCGGTGGTATTTATCTTCCATGGCT 241
 DB 185 TGAGATACCTGTCCAAAGTCCGGGAACTTATCCGGTGGTATTTATCTTCCATGGCT 242
 QY 242 TTTATCTTCCAACTTCTTCTTCTGATCTGATCTTCTTACACATCCGCTTCCATGGTCA 301
 DB 243 TTTATCTTCCAACTTCTTCTTCTGATCTGATCTTCTTACACATCCGCTTCCATGGTCA 302
 QY 302 TTTCTGTAGCCCAAGTTGTGCAATTTTCCCGCCGGAGGGCAAGTGGAGTGGACG 361
 DB 303 TTTCTGTAGCCCAAGTTGTGCAATTTTCCCGCCGGAGGGCAAGTGGAGTGGACG 360
 QY 362 ATGCTGGAAGTGTGATTAACCTGGGATCGGAAACCTCAAGCTCACTACCACTTCCG 421
 DB 361 ATGCTGGAAGTGTGATTAACCTGGGATCGGAAACCTCAAGCTCACTACCACTTCCG 420
 QY 422 TAAATGCTAATGGAATAATACCTCTGATCTGTGGCCACAGCCGCGTGGAAACCGGCT 481
 DB 421 TAAATGCTAATGGAATAATACCTCTGATCTGTGGCCACAGCCGCGTGGAAACCGGCT 480
 QY 482 TTGCGGTGGCTTGGCCATGCGGCACTTATGACCTTCCATCCATCGTTTTCAGCTTAA 541
 DB 481 TTGCGGTGGCTTGGCCATGCGGCACTTATGACCTTCCATCCATCGTTTTCAGCTTAA 540
 QY 542 TAGGAATGATCCAGTCCGAGGAACTCAAAATACATTAGAACGATCCGATCTTAA 601
 DB 541 TAGGAATGATCCAGTCCGAGGAACTCAAAATACATTAGAACGATCCGATCTTAA 600
 QY 602 CGTATTAACCGGAATCTTTCGAGCTGGAATACCGGTTGCAAGTGGTGGAAACCGGACTCG 661
 DB 601 CGTATTAACCGGAATCTTTCGAGCTGGAATACCGGTTGCAAGTGGTGGAAACCGGACTCG 660
 QY 662 GACCGAAGTGGAAACGATGATCCATCCATGCGGCAACGAGCTTAAACCATGAGAGT 721
 DB 661 GACCGAAGTGGAAACGATGATCCATCCATGCGGCAACGAGCTTAAACCATGAGAGT 720
 QY 722 TTTCAAAAGTGTGAGGAGGAGCAAGGACCATTTGCGGCTGCGGATTAACGACATATAG 781
 DB 721 TTTCAAAAGTGTGAGGAGGAGCAAGGACCATTTGCGGCTGCGGATTAACGACATATAG 780
 QY 782 ATATGTTGAGCGATGATTTGCGGCTTGTGAGGCTTGTATGCTGTAAGA 841
 DB 781 ATATGTTGAGCGATGATTTGCGGCTTGTGAGGCTTGTATGCTGTAAGA 839
 QY 842 ATGGGCAAGAAAGAAAGTCTGAGATGAGAGAGCTTTGTATGATGAGATTTGCTGCTTC 901
 DB 840 ATGGGCAAGAAAGAAAGTCTGAGATGAGAGAGCTTTGTATGATGAGATTTGCTGCTTC 899
 QY 902 TCAAGTATAGTTTGGGGTGAAGAAAGCGAGATTCGATTTGATTTGAAAGATCTTCCG 961
 DB 900 TCAAGTATAGTTTGGGGTGAAGAAAGCGAGATTCGATTTGATTTGAAAGATCTTCCG 959
 QY 962 TTTCTCCGCGCAAGCTTATCTTCACTGATGATGAAAGAGCTTCTGATCTTCTGCTCT 1021
 DB 960 TTTCTCCGCGCGCAAGCTTATCTTCACTGATGATGAAAGAGCTTCTGATCTTCTGCTCT 1019
 QY 1022 AGATTGTGTTATGATCTATTTATGAGAGGCTTGAATTTTGAAGAAACCTATCAAGT 1081
 DB 1020 AGATTGTGTTATGATCTATTTATGAGAGGCTTGAATTTTGAAGAAACCTATCAAGT 1079
 QY 1082 TTTCTAGCTCCAAAGCTATTTGTCATGCTCAAGTGGATGATTTTATTTATTAAC 1141
 DB 1080 TTTCTAGCTCCAAAGCTATTTGTCATGCTCAAGTGGATGATTTTATTTATTAAC 1139
 QY 1142 TCGATCAAAACATTTGT 1158

DB 1140 TCGATCAAAACATTTGT 1156
 RESULT 6
 AC024609
 LOCUS
 DEFINITION
 Arabidopsis thaliana chromosome I BAC F14P1 genomic sequence,
 complete sequence.
 AC024609.2 GI:7212002
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Arabidopsis thaliana (thale cress)
 Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
 1 (bases 1 to 90341)
 Federpiel,N.A., Palm,C.J., Conway,A.B., Conn,L., Hansen,N.F.,
 Altafi,H., Nguyen,M., Lam,B., Southwick,A., Bel,O., Buehler,E.,
 Chiu,C., Chlou,J., Choi,E., Dunn,P., Gonzalez,A., Hong,B., Kim,C.,
 Koo,T., Lee,J.M., Lenz,C., Li,J., Liu,A., Liu,K., Liu,S.,
 Mukharsky,N., Pham,P., Sakano,H., Schwartz,J., Shim,P.,
 Thaveri,A., Toriumi,M., Vayberg,M., Walker,M., Yu,G., Ecker,J.,
 Theologis,A. and Davis,R.W.
 Submitted (01-MAR-2000) DNA Sequencing and Technology Center,
 Stanford University, 855 California Avenue, Palo Alto, CA 94304,
 USA
 3 (bases 1 to 90341)
 Federpiel,N.A., Palm,C.J., Conway,A.B., Conn,L., Hansen,N.F.,
 Altafi,H., Nguyen,M., Lam,B., Southwick,A., Bel,O., Buehler,E.,
 Chiu,C., Chlou,J., Choi,E., Dunn,P., Gonzalez,A., Hong,B., Kim,C.,
 Koo,T., Lee,J.M., Lenz,C., Li,J., Liu,A., Liu,K., Liu,S.,
 Mukharsky,N., Pham,P., Sakano,H., Schwartz,J., Shim,P.,
 Thaveri,A., Toriumi,M., Vayberg,M., Walker,M., Yu,G., Ecker,J.,
 Theologis,A. and Davis,R.W.
 Direct Submission
 Submitted (09-MAR-2000) DNA Sequencing and Technology Center,
 Stanford University, 855 California Avenue, Palo Alto, CA 94304,
 USA
 4 (bases 1 to 90341)
 Federpiel,N.A., Palm,C.J., Conway,A.B., Conn,L., Hansen,N.F.,
 Altafi,H., Nguyen,M., Lam,B., Southwick,A., Ecker,J., Theologis,A.
 and Davis,R.W.
 Direct Submission
 Submitted (11-AUG-2000) DNA Sequencing and Technology Center,
 Stanford University, 855 California Avenue, Palo Alto, CA 94304,
 USA
 5 (bases 1 to 90341)
 Federpiel,N.A., Palm,C.J., Conway,A.B., Conn,L., Hansen,N.F.,
 Altafi,H., Nguyen,M., Lam,B., Southwick,A., Ecker,J., Theologis,A.
 and Davis,R.W.
 Direct Submission
 Submitted (12-SEP-2000) DNA Sequencing and Technology Center,
 Stanford University, 855 California Avenue, Palo Alto, CA 94304,
 USA
 COMMENT
 On Mar 9, 2000 this sequence version replaced gi:7121532.
 Bases 1-59,676 of IGF clone F14P1 overlap with IGF clone FEF9,
 gb|AC007797.
 e-mail for correspondence: arab@sequence.stanford.edu
 Genes with similarity to proteins in the databases are named
 'putative', '-like' or 'similar to'. Genes that have EST
 similarity but no significant protein similarity are described as

'unknown proteins'. Genes that are annotated based only on gene prediction software are described as 'hypothetical proteins'. The gene prediction programs used to predict genes include: Grai1 (Informatics Group, Oak Ridge National Laboratory, <http://compbio.ornl.gov/section/index.html>), GENSCAN (Chris Burge, <http://genomic.stanford.edu/~chris/GENSCAN.html>), and NeePlantGene (S.M. Hebsgaard, et al., CBS, Technical University of Denmark, <http://www.cbs.dtu.dk/NeePlantGene.html>).

FEATURES

source

1. 90341

/organism="Arabidopsis thaliana"

/mol_type="genomic DNA"

/db_xref="taxon:3702"

/chromosome="1"

/clone="F14P1"

1. 59676

/note="overlap with IGF clone F6F9, gb|AC007797, see Genbank record for BAC clone F6F9 for annotation in this region."

misc_feature

60210.. 61960

/gene="F14P1.1"

/join(60210..60507,60608..60778,60878..60927,61025..61081,61184..61249,61387..61461,61599..61717,61807..61960)

/gene="F14P1.1"

/note="Unknown protein"

/codon_start=1

/protein_id="AAF98407.1"

/db_xref="GI:9795589"

/translation="MRSIQAPVPCVPYRQGVALLVNCVSKEKTRSLRQKQFNGOTKNDKSOATVNLRLHLRYSKICLFSHSDGTSTABENEDDYKSVLEAVYKSGPDGMVVKMKDGRQLRCVNNPQCNLPYAPASAIYKMEGDTLLPIYLEMPSVLMAANTVQIARPTPTQYVKQMDMGREYVLRVTRVHEATYPAELYSKVGKSDCVSFDLRSDANINAVRCVPIQVKKYLAISGMKVITDSGKLSKQTPASDGLFTELDNRNGQPCFDTKEFDELVRNNQAVDEREDACKLSHPPLYVILVILNY"

63540.. 65756

/gene="F14P1.2"

/join(63540..63596,63685..63951,64046..64123,64205..64306,64387..64436,64527..64626,64741..64833,64927..65040,65139..65253,65359..65756)

/gene="F14P1.2"

/note="Hypothetical protein"

/codon_start=1

/protein_id="AAF98408.1"

/db_xref="GI:9795590"

/translation="MTNMTQMRDFTGDTITLEDPEPELDEVIRYPOGYGVKDEGRPVYIERLGVKADAKLQVTLERYLYHVEFEKTTVKEPACCIARHIDISSTI LDVQGLKNETKTARDLILOLQKIDSDNYPETLHMFIIINAGSGFLKLTGVSFIDPKYSKIHVLGNKQONKLEIMIDASQPDFFGCTCADQCGKSRSDKPMKQSEILKMGSGGTCRHAGAFSLSDQISSDKPTYSIKYSDTSAKSGSELEMAKSPKNTNKHVRLTPVSEYANINISPTVLSYEEVCPMYDKYVDVAMQLEMPNASEGPOTSSIGKIGSVRIHWSMLAFISFTLLASLALPOTKESQHLSSVRAELCDERIARESRPSPRSTTERTVITISSVLSRLDLEKQIENLHRSKSEMPHEKELINAAVYRVALBAE LITTKKV"

70742.. 73284

/gene="F14P1.3"

/join(70742..70816,70957..71382,72084..72173,72304..73284)

/gene="F14P1.3"

/note="Hypothetical protein"

/codon_start=1

/protein_id="AAF98406.1"

/db_xref="GI:9795588"

/translation="MEYMRVILHMKNGENSTYAKNSTAGSNIISLGRVDEAKKIMMSNERISIGIADLCSSGSPSLISINIVDTILNCPDLRVPBELRVSLNDLPNSD FNYICSLPERDYDNNKKEGLFGRGSGCFVSAVPGFYGRLFPRLHFWSSS SLHWISQKITGSHNRRLISWGTGSSSYHAGGIGIIEEEKIDIFNAPYAAASB ELKAVIEKESGFSIDRLISPIIDWEGSISSEYDLIVIRSKPELAKGRVSTIRAV VEPMLBPTGENVMDLPERYAKIVGEYFYVSSPRVAIVLISLVRAQ"

74552.. 76417

/gene="F14P1.4"

/join(74552..74694,74769..75090,75300..75449,75558..75735,75908..75987,76159..76245,76328..76417)

/gene="F14P1.4"

/note="Hypothetical protein"

/codon_start=1

/protein_id="AAF98407.1"

/db_xref="GI:9795589"

/translation="MEYMRVILHMKNGENSTYAKNSTAGSNIISLGRVDEAKKIMMSNERISIGIADLCSSGSPSLISINIVDTILNCPDLRVPBELRVSLNDLPNSD FNYICSLPERDYDNNKKEGLFGRGSGCFVSAVPGFYGRLFPRLHFWSSS SLHWISQKITGSHNRRLISWGTGSSSYHAGGIGIIEEEKIDIFNAPYAAASB ELKAVIEKESGFSIDRLISPIIDWEGSISSEYDLIVIRSKPELAKGRVSTIRAV VEPMLBPTGENVMDLPERYAKIVGEYFYVSSPRVAIVLISLVRAQ"

74552.. 76417

/gene="F14P1.4"

/join(74552..74694,74769..75090,75300..75449,75558..75735,75908..75987,76159..76245,76328..76417)

/gene="F14P1.4"

/note="Hypothetical protein"

/codon_start=1

/protein_id="AAF98407.1"

/db_xref="GI:9795589"

/translation="MEYMRVILHMKNGENSTYAKNSTAGSNIISLGRVDEAKKIMMSNERISIGIADLCSSGSPSLISINIVDTILNCPDLRVPBELRVSLNDLPNSD FNYICSLPERDYDNNKKEGLFGRGSGCFVSAVPGFYGRLFPRLHFWSSS SLHWISQKITGSHNRRLISWGTGSSSYHAGGIGIIEEEKIDIFNAPYAAASB ELKAVIEKESGFSIDRLISPIIDWEGSISSEYDLIVIRSKPELAKGRVSTIRAV VEPMLBPTGENVMDLPERYAKIVGEYFYVSSPRVAIVLISLVRAQ"

74552.. 76417

/gene="F14P1.4"

/join(74552..74694,74769..75090,75300..75449,75558..75735,75908..75987,76159..76245,76328..76417)

/gene="F14P1.4"

/note="Hypothetical protein"

/codon_start=1

/protein_id="AAF98400.1"

/db_xref="GI:9795582"

/translation="MFQDRKSTAGVPGSGDFPVIGETLQFMILVNSGKGFYEVRSRRIRYSGCFPTSLGETHVFISTESARVLVNDGWFTRYSIKSIGELVGDLSLCAPOHHNKLRSRLINLPEKRSSTALMVHEDLVYDAGGMEHRCGVVLLITLLOTPKAMCKRLVLEKEBEISQORVGVCEMLAFPLNLPTRFHKIMARGVMEDEKTLRERNENSHNNHEDFLQQLAVNDTLPOLTAELKDNLTLMVINGQDTASALTMVKYLGEMQKVLIDILIEOSQITTKASNKPLLEJDESEMYASKMVESLRMASVPMFPLVLQDCMEGIL"

complement (77375..78430)

/gene="F14P1.5"

complement (77375..78430)

/gene="F14P1.5"

/note="Unknown protein"

/codon_start=1

/protein_id="AAF98401.1"

/db_xref="GI:9795583"

/translation="MEAEKRAASMTTENLFOVKTYTGLSAVAFGICCLPIVYNQHLGFPIWDMFTLVKLNLFSLVSVSIIIFLIYLFKIQTEKPELDYDVAVAPPSTVTPEDSVNMSKREDDGYNNSTYRENNVAFQAVAVESYSYDDEVEMSSKSRRTSRSEKTEMRVVEYKTEETVETGTSWRSQMDGLSGEEFERPIITQFVSGOKMHIIRASGVDPQLENGYIHRASGVDDRVHVEHGAVPOMQNDVPOLONGGVPMWONGSVSOLQNSDVPOMONGVYLOLQNGDVPQWOTGVSQLONGVHEPQNICYQMKSRDQSGRRHRHGRSKRHRO"

80606..80776

/gene="F14P1.6"

80606..80776

/gene="F14P1.6"

/note="Similar to defensin, similar to defensin AMP1 (anti microbial protein)"

/codon_start=1

/protein_id="AAF98402.1"

/db_xref="GI:9795584"

/translation="MAVEGRICERSKTYTFCGNTRGCDQCKMERASHGACHAQFPGFACRCYVNC"

complement (81552..83365)

/gene="F14P1.7"

complement (81552..83365)

82329..82436,82543..82581,82670..82741,82820..83068,83225..83365)

/gene="F14P1.7"

/note="Unknown protein"

/codon_start=1

/protein_id="AAF98405.1"

/db_xref="GI:9795587"

/translation="MTAEALPKCEBAPLVYGLQPALINDVAVDMSLDPQIPGRDGSIVQDELEHMKELDAHSVAPLKKMAGSVTNTVRGIVSVFGVATIGVGDDEQGLPTPSNMGFSVSIIRLRKKGSTAOCTCYLDVDSGNTMRCLSAKYIQLADELSKEDTGSKMLVLRVAVNLQVLOAIIPAKQEGSVSLDLASFEMVNSKSELQDLES GNIDLCFANEDEAEVLELRQGEAPALAEFLRHRQMVAVTIGSGGCIADKQEVVHISAIGETVATDAGAGDLPASGFLYGLIKGLSLEECCKVSGSGSVIRALGGEVTPENQOMKQLOLKGFLPVPDIHN"

complement (85358..87513)

/gene="F14P1.8"

complement (85358..87513)

87001..87144,87399..87513)

/gene="F14P1.8"

/note="Unknown protein"

/codon_start=1

Query Match 85.2%; Score 1012; DB 15; Length 90341;

Best Local Similarity 88.5%; Pred. No. 8.6e-274;

Matches 1175; Conservative 0; Mismatches 0; Indels 153; Gaps 1;

QY 2 ATACATTAATCTCAACACAACTCTTAATTAATCTGATTAATACAAATGCGCGCATAG 61

DB 57480 ATACATTAATCTCAACACAACTCTTAATTAATCTGATTAATACAAATGCGCGCATAG 57539

QY 62 AGGACAGTCCAAAGCTTTTCTCTGTGTGTAATCTCGGCGGCTTTGAGATAGCAGGCTCC 121

DB 57540 AGGACAGTCCAAAGCTTTTCTCTGTGTGTAATCTCGGCGGCTTTGAGATAGCAGGCTCC 57599

Qy	122	CGACAAACGAGATACCGGTGATCCGCTGGAATAATGATTCACACGACCGCAAAACCGG	181		Db	58680	GGGGCTTGATATTTGGAAAAACCTATCAATGTTTCTAGCTCCAGCTACTATGTTC	58739
Db	57600	CGACAAACGAGATACCGGTGATCCGCTGGAATAATGATTCACACGACCGCAAAACCGG	57659		Qy	1109	ATGCTTAAGTTGATGATGTTATTTTATTAATCACTGCATCAAAACATTTGTTATGTTTA	1168
Qy	182	TGAGAAATCACTGTCCAAACAGTCCCGGAACCTTATCCGTCGTTTATTTCTTCATGCGT	241		Db	58740	ATGCTTAAGTTGATGATGTTATTTTATTAATCACTGCATCAAAACATTTGTTATGTTTA	58799
Db	57660	TGAGAAATCACTGTCCAAACAGTCCCGGAACCTTATCCGTCGTTTATTTCTTCATGCGT	57719		Qy	1169	CCCCAAA 1176	
Qy	242	TTTATCTTGGAATCTACTCTGCTGACGCTCTTAACACATCCGCTTCGATGCTTTTCA	301		Db	58800	CCCCAAA 58807	
Db	57720	TTTATCTTGGAATCTACTCTGCTGACGCTCTTAACACATCCGCTTCGATGCTTTTCA	57779		RESULT 7			
Qy	302	TTCTTGTAGCCCC	314		AC007797			
Db	57780	TTCTTGTAGCCCCACAGGTACATATTTTAATACTATACAAAAATATTAATGATCATTTAA	57839		LOCUS			
Qy	315	-----	314		DEFINITION			
Db	57840	AAGATAACTTACGGTATGTAAATATGTAACATGTAAAAAATAATGTACATTAATTGAGAT	57899		ACCESSION			
Qy	315	-----	328		VERSION			
Db	57900	TGATTTGTAAGTAAAGATATATAGTTAATTTTATTTTCTTGATATGACAGTTGTCAAT	57959		KEYWORDS			
Qy	329	TATTCGCCCGGGAGGGCAAGTGAAGTGAAGATGCTGAAGTGTGATTAACCTGGGCAT	388		SOURCE			
Db	57960	TATTCGCCCGGGAGGGCAAGTGAAGTGAAGATGCTGAAGTGTGATTAACCTGGGCAT	58019		ORGANISM			
Qy	388	CGGAAAACCTCAAAGTCACTCAACCACTTCGGTAAATGCTTAATGGAATAATACCTCAAC	448		REFERENCE			
Db	58020	CGGAAAACCTCAAAGTCACTCAACCACTTCGGTAAATGCTTAATGGAATAATACCTCAAC	58079		AUTHORS			
Qy	449	TGCTGGGGCCACAGCGCGGTGGGAAAAAGGGGCTTTGCGGTTGCGCTAGGCCATGCGGAA	508		JOURNAL			
Db	58080	TGCTGGGGCCACAGCGCGGTGGGAAAAAGGGGCTTTGCGGTTGCGCTAGGCCATGCGGAA	58139		REFERENCE			
Qy	509	CATTAGACCCATCATCACTGTTTCACTTAATAGAAATTTGATCCAGTCGACGAGAACTA	568		AUTHORS			
Db	58140	CATTAGACCCATCATCACTGTTTCACTTAATAGAAATTTGATCCAGTCGACGAGAACTA	58199		TITLE			
Qy	569	ACAAATTAATTGAACCGATCCGCATTTTAACTTAACGTAACCGGAATCTTTCAGCTGG	628		JOURNAL			
Db	58200	ACAAATTAATTGAACCGATCCGCATTTTAACTTAACGTAACCGGAATCTTTCAGCTGG	58259		AUTHORS			
Qy	629	ACATACCGGTTGAGTGTGGGAGACCGGACCTCGACCGGAGTGAACAACGTAATGCCAC	688		REFERENCE			
Db	58260	ACATACCGGTTGAGTGTGGGAGACCGGACCTCGACCGGAGTGAACAACGTAATGCCAC	58319		AUTHORS			
Qy	689	CATGGCACAACGACCTTAAACATAGAGAGTTTAAACAAAGAGTGAAGCGACGAAGAAG	748		TITLE			
Db	58320	CATGGCACAACGACCTTAAACATAGAGAGTTTAAACAAAGAGTGAAGCGACGAAGAAG	58379		JOURNAL			
Qy	749	CCCATTTTGGTGGCTGCGATTACGACATATGGAATGTTGGACATGATTTTCCCGGTT	808		AUTHORS			
Db	58380	CCCATTTTGGTGGCTGCGATTACGACATATGGAATGTTGGACATGATTTTCCCGGTT	58439		TITLE			
Qy	809	TTGTTGGGTTTATGAGCCGTTGATGTATGTAAGATGGGCAAAAGAAAAAGCTGAGATGA	868		JOURNAL			
Db	58440	TTGTTGGGTTTATGAGCCGTTGATGTATGTAAGATGGGCAAAAGAAAAAGCTGAGATGA	58499		AUTHORS			
Qy	869	GGAGCTTTGTAGTGAATGTGTGTGCGTTTCTCAAGTATGATTGTGTGGGTGAAAAAG	928		REFERENCE			
Db	58500	GGAGCTTTGTAGTGAATGTGTGTGCGTTTCTCAAGTATGATTGTGTGGGTGAAAAAG	58559		AUTHORS			
Qy	929	CGGAATTCGATTTGTTGAAGATCCTTCCGTTTCTCCCGGCAAGCTTGATCCTTCAAC	988		TITLE			
Db	58560	CGGAATTCGATTTGTTGAAGATCCTTCCGTTTCTCCCGGCAAGCTTGATCCTTCAAC	58619		JOURNAL			
Qy	989	CTGAGTTGGAAGAACCTTCTGATCTTGATCTGTAGATTTGTTGTTATGTAATTAATCA	1048		COMMENT			
Db	58620	CTGAGTTGGAAGAACCTTCTGATCTTGATCTGTAGATTTGTTGTTATGTAATTAATCA	58679		FEATURES			
Qy	1049	GGGGCTTGAATATTTGAAAAACCTATCAATGTTTCTAGCTCAAGCTAGCTATTTGTTTC	1108		SOURCE			

gene
/chromosome="1"
/clone="F6f9"
2448. .4393
/gene="F6f9.1"
join(2448. .2640, 2712. .2995, 3068. .3502, 3589. .3727,
3808. .3974, 4064. .4393)
/gene="F6f9.1"
/note="Similar to endo-beta-1,4-glucanase"
/codon_start=1
/protein_id="AAG12562.1"
/db_xref="GI:10086502"
/translation="MVAKPRSCCCSVPIGVIITLAIITAVIFETIRHNSHDDDS
MYKYNALAKIAMQFPDIQKSGLENNESIKRSGSLKDSSEASIDSKGLYADGDM
KFGPPAFATVLSWSTLEIGDQMASLNLDDHADSILKWTDFLINHSPNVLYIQV
GDPTDHCWDRPETWTKRKLTKIDIKTEGETEVAALTAAMAASLVFKESDITKSS
TLKHAQLPDPADNNRGYSVNIPEVQSYNSYSGDELLMAASMLYHATBEDQTYLD
FVSENGEFGNFGSGFWFMDNKLPGTHILLSRLTFPKGLSGSGKQGGKEETAEAM
CGIIPSPATSSRTDGLIIVSEEMNLQHPVSSAFIATLTSOYMLTSYKELSCSDQ
SPKSDLRKPARSQADYMLGKNPEKSYLYGKPEFPHHKGASIPDAITGCKDQ
PKIINSDEPNPNAVIGALVGSPINDTFIDRNNMQNEPSTINSALVGLSLVLT
SSSVESF"
/gene="F6f9.2"
6059. .7881
join(6059. .6600, 6686. .7091, 7183. .7443, 7547. .7633,
7747. .7881)
/gene="F6f9.2"
/product="guilfate adenylyltransferase"
/codon_start=1
/protein_id="AAG12541.1"
/db_xref="GI:10086481"
/translation="MSLMIRSSVYSHITLQPRNSKSSFTNQISPLSSNNPILNL
VYKRNLTMOVSMTKSLIDPDGGELVEIIPETIGVAKASETMPVAKNLQIDL
EWAHVISGEWASPLKGFMRDEYIQSHFNSLRKNGTFVMSLPIVLAIDDDTKED
IGSSBNVALVPOGDIIGSLRSVEIYKNEKERIARTGITSPLPYVEEYITSGNWL
IGGDLVEFEBIKYNDGDHYRLSPKQREEDNRQADAVAFQLRNPHNGHALNMD
TRKRLMEYKXNPILLHPGLGFTKADVDVDMEOHSEVLEDVGLDPTKTTIVSIFP
SPMYAGPTEVOMAKARINAGANFYIVGRDPAQMGHPTEKDLQDDHCKRISNAP
GLEKTLNLPRIYVAIYDTIEKKMAFPDSSRAKEFLISGTMRRTARTGENPDPGFMC
SGMWLVKYESIQSEAKQAVSA"
complement(8394. .9399)
/gene="F6f9.3"
complement(join(8394. .8526, 8930. .9215, 9321. .9399))
/gene="F6f9.3"
/codon_start=1
/product="vacuolar H+-pumping ATPase"
/protein_id="AAG12542.1"
/db_xref="GI:10086482"
/translation="MASTFGSDTAPPPFGFLGAALVFSOMGAAYGTAKSGVASM
GWRPELVMSIVPVVMAGVLIGYGLINIVTISGINPRKASYLLPFGYHLSGLAC
GLAGLSAGMAIGIVDGRANMQQPLFGMTILILFAELALVGLIVGLIILSSRAG
QSRAE"
14164. .15678
/gene="F6f9.4"
14164. .15678
/gene="F6f9.4"
/note="Unknown Protein"
/codon_start=1
/protein_id="AAG12543.1"
/db_xref="GI:10086483"
/translation="MHQMLHNRVVWYDRTFNGPSNISLPNGCRNDPODAVKIDC
TASISDVATNTRILPTVQSNTRKSSGSGVPRGVNCGSDRDGELKTPTRSCNN
QCDVEMKNGLKRRRYASNHILPDKQIYMGQGOQNTFFPRTTNPNVALPFLAE
THQGOENLNLPEVPMATDNLFWFANRAIILIDYGVNTVKTPEPALPGDPPNYBST
GSAVLPLKNI.EADVNTGLVCGGAKSGYNLARKKTPKALDTCARIKINDAKEM
AVEKMPHARVMSAIPLPNDGVLLINGSGFATAMELGRTPVLAJDLYHBNVGSRF
ESLRPTTIPMYGSAIILLRDRVLVINGSNPHAFYNTYGLPFTBELSLAEFNVYOR
EBSNLARKIISPEBOSMIKYGTNLKIKESYGEVTPAKYTMLEPPTTTHSPMNGRV
LVLDNVKFTKSGSPMIEVORTPRSNINMPGYMIFVYNQDIPSEGVWAKLE"
complement(116924. .17682)
/gene="F6f9.5"
complement(join(16924. .17094, 17228. .17318, 17417. .17496,
17611. .17682))
/gene="F6f9.5"

gene
/codon_start=1
/product="putative histone H3"
/protein_id="AAG12563.1"
/db_xref="GI:10086503"
/translation="MARTKOTAKNSYGGKPRKELATKARKTRRPYRGVKKRHRP
PGVVALREIKYOKSTDLIRKLPRQRLVREIQADPKVDLRFGSHAVLALQEAAEVY
VGLPEDTNCAIHAKEVTTMSKDIOLRKIRGBRA"
18898. .22068
/gene="F6f9.7"
join(18898. .19093, 19182. .19243, 19336. .19403, 19510. .19631,
19701. .19824, 19937. .19988, 20276. .20393, 20887. .21016,
21098. .21220, 21301. .21354, 21480. .21661, 21734. .22068)
/gene="F6f9.7"
/note="Unknown Protein"
/codon_start=1
/protein_id="AAG12544.1"
/db_xref="GI:10086484"
/translation="MAEAMNSPKKDSSEKNGELLFGCATAMDILGKRGAMEGLT
VSEFTRRLPVGNIRFVATGCASTFHCVALDVBGRCTTGWNERGQGLGHDMITDRDPT
VSGLSRHKIKVKAAGRNTVVVSDGSLGFGMNRYGQGLGASANGFVSVESSTP
LPCVSDDEVTVNACGADFTVWLSTEGASILFAGLPQYQLGHTDNERNKDSVRL
AYEAPQRPKALIASIAGETIVKACGTHVTFWYIKTKYNLFLSGGGQLYMKGXIKN
NGDDMYPKPMMDLSGMNLRMDSGSMHPVAGDSICIEMHAYGELGYPGGOSS
AARKYDMLEGHWMTGTDIADRLFOYTKAREPFRILALRKIKRYTSSRRGTSSTSL
HCLICQLLEYDYGKSLBAESVEEKEETLAPQOAVTYTKGASKKAKASASDSEOD
SDEINDKEKEYGSDADSDYSDEGEANGKQASARGRGARGRGRTSNGKAP
VKTGRRGRPRKS"
24177. .27256
/gene="F6f9.8"
join(24177. .24269, 24579. .25172, 25331. .25510, 25592. .25651,
25766. .27256)
/gene="F6f9.8"
/note="Unknown Protein"
/codon_start=1
/protein_id="AAG12559.1"
/db_xref="GI:10086449"
/translation="MGRSPASSCTLRISCSGGDDTTSADNRSTALENKSSGDRKGNFR
KSGKORGLITVSVSTTPASRTRETTESLALSRSPPDNNSVSEKQDSFVDEKSO
LPLVTVVABPVDEKQTQSVIEEKTLLSVEEQIDHTEVSYPVSEDEKVEDKDESDIG
TELOGSADDAKIEEDVTSEVMASKVSDVDVIVRSEDEKVEDKDESDIG
VVIQAAVRGFLARBELLRSKVYKIQAAVRGHLVRSQAMGSLRCVOAIVMOMVBAR
HSTKQDSRYATSDSKSEPMNAAQKLENNKPAHLMSESTKTPINIKCPTKPSAMN
WLERMNSVPRPETSANLTTEQNLIEFQNKIISQADPANSSTVENEKTEIDMSY
EASKVQNVLESTETKMSQYDSPEASAVYTDISQDPLAAPDSLSEPEYDQI
KSLARKVSNPSEFIAQSKFELTSSGKMKMTQSKDVPVAGEBKTDIDSDTNT
IKDHSLEDVTPALISGSECGTELSYSLDTEKSDAEGAERPAKLEDDTPTFD
QAELEIDVDKATSMGTVDPKKEVENADEVEISATHEPVISTPDSKKRAEDSG
POAVYLSKALPMTTIESOATPAQASASSVARKKSKSSGSKRYSKITTSXP
QELGTEBALDEBGRKQSKGRITSFETDDEABESSGKNSLPRFMPTOSAKAKYDQH
NSPRSSPDLOERDVVSAKRHSILPGVTNKOVSPIRORSASQAOGTGKNLLOSFL
CHNSQF"
28911. .31062
/gene="F6f9.9"
join(28911. .28985, 29480. .29677, 29976. .30100, 30198. .31062)
/gene="F6f9.9"
/note="Hypothetical Protein"
/codon_start=1
/protein_id="AAG12545.1"
/db_xref="GI:10086485"
/translation="MOENAPRTMSQDVQVLAISNKVFDVNRLLFISRSPQSGSESOD
HIOAKSPPLASHSDDLPLPGRGPRPANSQIKLSIPILKIKKCSQOILLDEWKYVA
GDSKVEAKQNERELNLEAFYFGASITPPNSVPAVDVDSHDDQITVILPLVED
DIDLAMSADFPQSGVDVGTESITDENTSTSTSLPAPDIPAAALASIASNSKEQSM
IIODDLILKILSNPKLEVNIVARGSGASNTSSVSSSTHEANVATASISNSGO
FYAOPRTTIPMAVYPAPOQNPYGPAPADASYKYLIOOHGGDROETPEVQIAG
YRNLOPGGPDVEMVNSNNORPDSDPKIMKCMYNSARGCHGANCMYQHDAT
PYQPRILNNGNINTSMQAKXMRPFRD"
complement(31753. .35330)
Query Match 85.2%; Score 1012; DB 15; Length 119942;
Best Local Similarity 98.5%; Pred. No. 8,6e-274;
Matches 1175; Conservative 0; Mismatches 0; Indels 153; Gaps 1;

QY	2	TATCACTAATCTTCAACACAACTTTAATTAATCACTGTTAAATACAAATAGCGCGCATAG	61
Db	117746	ATACATAAATCTTCAACACAACTTTAATTAATTAATCACTGTTAAATACAAATAGCGCGCATAG	117805
QY	62	AGACAGTCTCAACGTTTTCTCTGTGTGTAACTCCGCGGCTTTTGAGATAGGAGCCTCC	121
Db	117806	AGGACAGTCCAAAGTTTTCTCTGTGTGTAACTCCGCGGCTTTTGAGATAGGAGCCTCC	117865
QY	122	CGACAAACGAGATACCGGTGGATCCGGTGGAAAATGATTCACACGACCGCCAAAACCG	181
Db	117866	CGACAAACGAGATACCGGTGGATCCGGTGGAAAATGATTCACACGACCGCCAAAACCG	117925
QY	182	TSAGAAATCAACGTGCCAAGTCGCCGAACTTATCCGCTGTTTATCTTCCATGGCT	241
Db	117926	TSAGAAATCAACGTGCCAAGTCGCCGAACTTATCCGCTGTTTATCTTCCATGGCT	117985
QY	242	TTTATCTTGGCAACTACTCTACTCTGAGGTTCTTAACACATCGCTTGGCATGGTTACA	301
Db	117986	TTTATCTTGGCAACTACTCTACTCTGAGGTTCTTAACACATCGCTTGGCATGGTTACA	118045
QY	302	TTCTTGTAGCCCC-----	314
Db	118046	TTCTTGTAGCCCCACAGTACATTAATTTACTATACAAAATAATATGAATCATTTAA	118105
QY	315	-----	314
Db	118106	AAGATTAACCTTACGGTATGTAAATGTAAATATGTAAACAAATGTACTTAATTGAGAT	118165
QY	315	-----ACAGTTGTGCAAT	328
Db	118166	TGATTTGTTACGTGAAAAGATTAATTAATTTTATATCTTGAAATGACAGTTGTGCAAT	118225
QY	329	TATTCCTCCGCGGAGGAGGCAATGTGAATGTGACATGCTGGAAGTGTATTAACCTGGGCAT	388
Db	118226	TATTCCTCCGCGGAGGAGGCAATGTGAATGTGACATGCTGGAAGTGTATTAACCTGGGCAT	118285
QY	389	CGGAAAACTCTCAAGCTCACTTAACCACTTCGTTAATGTCTAATGGAAAAATACACTCAC	448
Db	118286	CGGAAAACTCTCAAGCTCACTTAACCACTTCGTTAATGTCTAATGGAAAAATACACTCAC	118345
QY	449	TGCTGGGCGCACAGCGCGGTGGGAAAAAGCGCGTTGCGGTGGCGCTAAGGCATGCGCGCA	508
Db	118346	TGCTGGGCGCACAGCGCGGTGGGAAAAAGCGCGTTGCGGTGGCGCTAAGGCATGCGCGCA	118405
QY	509	CATTAGACCCCATTCATCAAGTTTTCACTTAATAGAAATGTATCCAGTCGACGAACTTA	568
Db	118406	CATTAGACCCCATTCATCAAGTTTTCACTTAATAGAAATGTATCCAGTCGACGAACTTA	118465
QY	569	ACAAATACATTAGAAACCGATCCGCATATCTTAAACGATTAACCGGAATCTTTGAGCTGG	628
Db	118466	ACAAATACATTAGAAACCGATCCGCATATCTTAAACGATTAACCGGAATCTTTGAGCTGG	118525
QY	629	ACATACCGGTTGACAGTGTGGGGAACCGGACTCGGACCGAAGTGGAAACATCGTATCCAC	688
Db	118526	ACATACCGGTTGACAGTGTGGGGAACCGGACTCGGACCGAAGTGGAAACATCGTATCCAC	118585
QY	689	CATGCGCAACCAACGACTTAAACCATGAGAGTTTTTACAAAGTGTAAAGCGACCAAAAG	748
Db	118586	CATGCGCAACCAACGACTTAAACCATGAGAGTTTTTACAAAGTGTAAAGCGACCAAAAG	118645
QY	749	CCCATTTTGTGGCTGCGGATTAACGACATATGATATGTTGGACGATATTTGCCGGTT	808
Db	118646	CCCATTTTGTGGCTGCGGATTAACGACATATGATATGTTGGACGATATTTGCCGGTT	118705
QY	809	TTGTTGGGTTTATGCGCGGTTGATATGTATGAATGGGCAAAAGAAAGTCTGAGATGA	868
Db	118706	TTGTTGGGTTTATGCGCGGTTGATATGTATGAATGGGCAAAAGAAAGTCTGAGATGA	118765
QY	869	GGAGCTTTGTAGTGTGAATTTGTGTTGCTTTCTCAAGTATAGTTTGTGGGCTGAAAAAG	928
Db	118766	GGAGCTTTGTAGTGTGAATTTGTGTTGCTTTCTCAAGTATAGTTTGTGGGCTGAAAAAG	118825
QY	929	CGGAGATTCGATTTGATTTGAAAGATCTTCCGTTTTCTCCGCGCAAGCTTGATCTTTCAC	988

QY	Db	118826	CGAAGTTCGATTGATTGGAAGAGTCCCTCGCTTCCGGCAAGCTTGATCTTCAC	118885
QY		989	CTGAGTTGGAGAGAGCTTCGTGATCTTCCTGCTAGATTTGTGTATGTAATATACAGA	1048
Db		118886	CTGAGTTGGAGAGAGCTTCGTGATCTTCCTGCTAGATTTGTGTATGTAATATACAGA	118945
QY		1049	GGGGTCTTGAAATATTTGAAAAAAGCTATCAATGTTTCTAGCTCCAGCTAGATTTGTTT	1108
Db		118946	GGGGTCTTGAAATATTTGAAAAAAGCTATCAATGTTTCTAGCTCCAGCTAGATTTGTTT	119005
QY		1109	ATGTCCTAGTTCGATGATGATATTTTATTAATCTGATCAAAACATTTGTATAGTTTAA	1168
Db		119006	ATGTCCTAGTTCGATGATGATATTTTATTAATCTGATCAAAACATTTGTATAGTTTAA	119065
QY		1169	CCCCAAA 1176	
Db		119066	CCCCAAA 119073	
RESULT 8				
AX412268				
LOCUS		AX412268	975 bp	DNA
DEFINITION		Sequence 32 from Patent WO0222675.		linear
ACCESSION		AX412268		
VERSION		AX412268.1		GI:21444726
KEYWORDS				
SOURCE				
ORGANISM				
		Arabidopsis thaliana (thale cress)		
		Arabidopsis thaliana		
		Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Eudicotyledons; core eudicotyledons; Rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.		
REFERENCE				
AUTHORS		1		
TITLE		Glazebrook, J., Wang, X., Dangl, J. L., Bulgem, T., and Zhu, T.		
JOURNAL		Plant gene, the expression of which are altered by pathogen infection		
		Patent: WO 0222675-A 32 21-MAR-2002;		
		Shengela Participations AG (CH) ; UNIVERSITY OF NORTH CAROLINA AT CHAPEL HILL (US) ; Glazebrook, Jan (US) ; Wang, Xun (US) ; Dangl, Jeffrey L. (US) ; Bulgem, Thomas (US)		
FEATURES				
source		location/Qualifiers		
		1..975		
		/organism="Arabidopsis thaliana"		
		/mol_type="unassigned DNA"		
		/db_xref="taxon:3702"		
ORIGIN				
Query Match		82.1%;	Score 975;	DB 6;
Best Local Similarity		100.0%;	Fred. No. 2.3e-263;	Length 975;
Matches		975;	Conservative 0;	Mismatches 0;
			Indels 0;	Gaps 0;
QY		49	ATGCGCGCATAGAGAGACAGTCCACAGTTCCTCTGTGTAACTCCGCGCTTTTGAG	108
Db		1	ATGCGCGCATAGAGAGACAGTCCACAGTTCCTCTGTGTAACTCCGCGCTTTTGAG	60
QY		109	ATAGGACGCTCCCGACCAACCGAGATACCGGTGATCCGGTGGAAAAATGATTAACAGCA	168
Db		61	ATAGGACGCTCCCGACCAACCGAGATACCGGTGATCCGGTGGAAAAATGATTAACAGCA	120
QY		169	CGGCGAAAAACCGGTGAGATCACTGTGCCAACAGTCCGCGAACTTATCCGTCGTTTAA	228
Db		121	CGGCGAAAAACCGGTGAGATCACTGTGCCAACAGTCCGCGAACTTATCCGTCGTTTAA	180
QY		229	TTCCTTCATAGGCTTTTATCTTGCAACATCTCTGATGATCTTGAACCATCGCT	288
Db		181	TTCCTTCATAGGCTTTTATCTTGCAACATCTCTGATGATCTTGAACCATCGCT	240
QY		289	TGCAATGTTATCATCTTGTAGCCCCACAGTTGTGCAATTATTCGCGCGGAGGGCAA	348
Db		241	TGCAATGTTATCATCTTGTAGCCCCACAGTTGTGCAATTATTCGCGCGGAGGGCAA	300
QY		349	GTGGAAGTGACGATGCTGGAAGTGTATTAATCTGGGCAATCGGAAACTCAAGCTCAC	408

Db 301 GTGAGTGTGACGATGCTGGAAGTGTGATTAACCTGGGCAATCGGAAACCTCAAGCTCAC 360
Qy 409 CTACCAACTTGGTAAATGCTAATGAAATAACCTCACTGCTGGCCACAGCCGCT 468
Db 361 CTACCAACTTGGTAAATGCTAATGAAATAACCTCACTGCTGGCCACAGCCGCT 420
Qy 469 GGGAAAACGGGCTTTGGGGTTGGGCTAGGGCCATGCGCAACATTAGACCCATCCATACG 528
Db 421 GGGAAAACGGGCTTTGGGGTTGGGCTAGGGCCATGCGCAACATTAGACCCATCCATACG 480
Qy 529 TTTTCACTCTAATAGAAATTTGATCGAGTCAGAACTAACTAATCATTAGAACCGAT 588
Db 481 TTTTCACTCTAATAGAAATTTGATCGAGTCAGAACTAACTAATCATTAGAACCGAT 540
Qy 589 CCGCATATCTTAACGTAATTAACCGGAATCTTTGAGCTGACATACCGGTTGCACTG 648
Db 541 CCGCATATCTTAACGTAATTAACCGGAATCTTTGAGCTGACATACCGGTTGCACTG 600
Qy 649 GGAACCGGACTCGGACCGAAGTGGACAAAGTGATGCAACATGCGACCAACGAGCTTA 708
Db 601 GGAACCGGACTCGGACCGAAGTGGACAAAGTGATGCAACATGCGACCAACGAGCTTA 660
Qy 709 AACCATGAGAGATTTTTCAAAAGAGTGAAGCGACGAAGCCATTTCGTGCTGCGGAT 768
Db 661 AACCATGAGAGATTTTTCAAAAGAGTGAAGCGACGAAGCCATTTCGTGCTGCGGAT 720
Qy 769 TACGACATATGATGATGTTGAGACGATGTTGCCGGTTTGTGGGTTTATGCGCGT 828
Db 721 TACGACATATGATGATGTTGAGACGATGTTGCCGGTTTGTGGGTTTATGCGCGT 780
Qy 829 TGTATGTTAGAAATGGGCAAGAAAGAAAGCTGAGATGAGAGCTTTGATAGTGAAT 888
Db 781 TGTATGTTAGAAATGGGCAAGAAAGAAAGCTGAGATGAGAGCTTTGATAGTGAAT 840
Qy 889 GTGGTTCGTTTCTCAAGTATGTTGTGGGGTGAAGAAAGCGAGATTCATGATTTG 948
Db 841 GTGGTTCGTTTCTCAAGTATGTTGTGGGGTGAAGAAAGCGAGATTCATGATTTG 900
Qy 949 AAGATATCTTCGTTTCTCGGCGCAAGCTTGATCTTCACCTGATGGAAGAAAGCTTCT 1008
Db 901 AAGATATCTTCGTTTCTCGGCGCAAGCTTGATCTTCACCTGATGGAAGAAAGCTTCT 960
Qy 1009 GGATCTTCGCTAG 1023
Db 961 GGATCTTCGCTAG 975

RESULT 9
AX412469 975 bp DNA linear PAT 14-JUN-2002
LOCUS AX412469
DEFINITION Sequence 233 from Patent WO0222675.
ACCESSION AX412469
VERSION AX412469.1 GI:21444927
KEYWORDS
SOURCE
ORGANISM Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1
Glazebrook J., Wang X., Dangl J.L., Bulgem T. and Zhu T.
Plant genes, the expression of which are altered by pathogen
infection
Patent: WO 0222675-A 233 21-MAR-2002;
Syngenta Participations AG (CH) ; UNIVERSITY OF NORTH CAROLINA AT
CHAPEL HILL (US) ; Glazebrook, Jan (US) ; Wang, Xun (US) ; Dangl,
Jeffrey L. (US) ; Bulgem, Thomas (US)
FEATURES
Source
1..975
/organism="Arabidopsis thaliana"
/mol_type="unassigned DNA"
/db_xref="taxon:3702"

ORIGIN

Query Match 82.1%; Score 975; DB 6; Length 975;
Best Local Similarity 100.0%; Pred. No. 2.3e-263;
Matches 975; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 49 ATGGCGGCGATTAAGGACAGTCCAAAGTTTCTCTGTGGTAACTCCGGCGGCTTTTGA 108
Db 1 ATGGCGGCGATTAAGGACAGTCCAAAGTTTCTCTGTGGTAACTCCGGCGGCTTTTGA 60
Qy 109 ATAGGAGCTCCCGCAACCGAGATACCGGTGATCCGGTGGAAATGATTTCAACAGCA 168
Db 61 ATAGGAGCTCCCGCAACCGAGATACCGGTGATCCGGTGGAAATGATTTCAACAGCA 120
Qy 169 CCGGCAAAACCGGTGAGATCACCTGTCCAAAGTGGCCGGAACCTTATCCGCTGTTTA 228
Db 121 CCGGCAAAACCGGTGAGATCACCTGTCCAAAGTGGCCGGAACCTTATCCGCTGTTTA 180
Qy 229 TTCTTCATAGGCTTTTATCTTGCAACTACTTCTAGAGTTCTTAACTCAATCGCT 288
Db 181 TTCTTCATAGGCTTTTATCTTGCAACTACTTCTAGAGTTCTTAACTCAATCGCT 240
Qy 289 TCGCATGTTTACATTTCTTGTAGCCCAAGTTGTGAAATTATGCGCGCGGAGGCA 348
Db 241 TCGCATGTTTACATTTCTTGTAGCCCAAGTTGTGAAATTATGCGCGCGGAGGCA 300
Qy 349 GTGAAAGTGAAGATGCTGGAATGTGATTAATGCGGCATCGGAAACCTCAAGCTCAC 408
Db 301 GTGAAAGTGAAGATGCTGGAATGTGATTAATGCGGCATCGGAAACCTCAAGCTCAC 360
Qy 409 CTACCAACTTGGTAAATGCTAATGAAATAACCTCACTGCTGGGCCACAGCCGCT 468
Db 361 CTACCAACTTGGTAAATGCTAATGAAATAACCTCACTGCTGGGCCACAGCCGCT 420
Qy 469 GGGAAAACGGGCTTTGGGGTTGGGCTAGGGCCATGCGCAACATTAGACCCATCCATACG 528
Db 421 GGGAAAACGGGCTTTGGGGTTGGGCTAGGGCCATGCGCAACATTAGACCCATCCATACG 480
Qy 529 TTTTCACTCTAATAGAAATTTGATCGAGTCAGAACTAACTAATCATTAGAACCGAT 588
Db 481 TTTTCACTCTAATAGAAATTTGATCGAGTCAGAACTAACTAATCATTAGAACCGAT 540
Qy 589 CCGCATATCTTAACGTAATTAACCGGAATCTTTGAGCTGAGACATACCGGTTGAGTGG 648
Db 541 CCGCATATCTTAACGTAATTAACCGGAATCTTTGAGCTGAGACATACCGGTTGAGTGG 600
Qy 649 GGAACCGGACTCGGACCGAAGTGGACAAAGTGATGCAACATGCGACCAACGAGCTTA 708
Db 601 GGAACCGGACTCGGACCGAAGTGGACAAAGTGATGCAACATGCGACCAACGAGCTTA 660
Qy 709 AACCATGAGAGATTTTTCAAAAGAGTGAAGCGACGAAGCCATTTCGTGCTGCGGAT 768
Db 661 AACCATGAGAGATTTTTCAAAAGAGTGAAGCGACGAAGCCATTTCGTGCTGCGGAT 720
Qy 769 TACGACATATGATGATGTTGAGACGATGTTGCCGGTTTGTGGGTTTATGCGCGT 828
Db 721 TACGACATATGATGATGTTGAGACGATGTTGCCGGTTTGTGGGTTTATGCGCGT 780
Qy 829 TGTATGTTAGAAATGGGCAAGAAAGAAAGCTGAGATGAGAGCTTTGATGTTGAAT 888
Db 781 TGTATGTTAGAAATGGGCAAGAAAGAAAGCTGAGATGAGAGCTTTGATGTTGAAT 840
Qy 889 GTGGTTCGTTTCTCAAGTATGTTGTGGGGTGAAGAAAGCGAGATTCATGATTTG 948
Db 841 GTGGTTCGTTTCTCAAGTATGTTGTGGGGTGAAGAAAGCGAGATTCATGATTTG 900
Qy 949 AAGATATCTTCGTTTCTCGGCGCAAGCTTGATCTTCACCTGATGGAAGAAAGCTTCT 1008
Db 901 AAGATATCTTCGTTTCTCGGCGCAAGCTTGATCTTCACCTGATGGAAGAAAGCTTCT 960
Qy 1009 GGATCTTCGCTAG 1023
Db 961 GGATCTTCGCTAG 975

RESULT 10	AX505463	975 bp	DNA	linear	PAT 27-SEP-2002
LOCUS	AX505463				
DEFINITION	Sequence 158 from Patent WO216655.				
ACCESSION	AX505463				
VERSION	AX505463.1	GI:23386700			
KEYWORDS					
SOURCE					
ORGANISM	Arabidopsis thaliana (thale cress)				
REFERENCE	1				
AUTHORS	Harper, J.F., Kreps, J., Wang, X. and Zhu, T.				
TITLE	Stress-regulated genes of plants, transgenic plants containing same, and methods of use				
JOURNAL	Patent: WO 0216655-A 158 28-FEB-2002;				
	The Scripps Research Institute (US) ; Syngenta Participations AG (CH)				
FEATURES					
source	Location/Qualifiers				
	1..975				
	/organism="Arabidopsis thaliana"				
	/mol_type="unassigned DNA"				
	/db_xref="taxon:3702"				
ORIGIN					
Query Match	82.1%;	Score 975;	DB 6;	Length 975;	
Best Local Similarity	100.0%;	Pred. No. 2,36-263;			
Matches 975;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;	
QY	49	ATGGCGGAGATGAGAGACAGTCCACGTTTCTCTGATGATTAACCCGGCGCTTTGAG	108		
DB	1	ATGGCGGAGATGAGAGACAGTCCACGTTTCTCTCTGATGATTAACCCGGCGCTTTGAG	60		
QY	109	ATAGGACGCTCCCGACAAACCGAGATACCGGTGATCCGGTGAATAATGATTCAACAGCA	168		
DB	61	ATAGGACGCTCCCGACAAACCGAGATACCGGTGATCCGGTGAATAATGATTCAACAGCA	120		
QY	169	CCGCCAAAACCGGTGAGATATCCTGTGCAACAGTCCCGGAACTTATCCCGTGTTTA	228		
DB	121	CCGCCAAAACCGGTGAGATATCCTGTGCAACAGTCCCGGAACTTATCCCGTGTTTA	180		
QY	229	TTCTTCCAAATGGTTTATCTTGTGCAACACTCTTCACTGACGTTCTTAACCAACATCGCT	288		
DB	181	TTCTTCCAAATGGTTTATCTTGTGCAACACTCTTCACTGACGTTCTTAACCAACATCGCT	240		
QY	289	TCGATGAGTTACATTTCTGTAGACCCCAAGTTGTCAAAATTATTCGCCGGGAGGCA	348		
DB	241	TCGATGAGTTACATTTCTGTAGACCCCAAGTTGTCAAAATTATTCGCCGGGAGGCA	300		
QY	349	GTGGAAGTGAAGCAAGTCTGGAAGTGTGATTAACCTGGGATCGGAAAACCTCAAACCTAC	408		
DB	301	GTGGAAGTGAAGCAAGTCTGGAAGTGTGATTAACCTGGGATCGGAAAACCTCAAACCTAC	360		
QY	409	CTACCAACTTGGTGAATGCTAATGGAATAATCACCTCACTGTTGGGCAACAGCGCGAT	468		
DB	361	CTACCAACTTGGTGAATGCTAATGGAATAATCACCTCACTGTTGGGCAACAGCGCGAT	420		
QY	469	GGGAAAACGGCGTTTGGCGTTGCGCTAAGCCATGCGCAACCTTAAGACCCATCATACG	528		
DB	421	GGGAAAACGGCGTTTGGCGTTGCGCTAAGCCATGCGCAACCTTAAGACCCATCATACG	480		
QY	529	TTTTCAGCTTAATGGAATTGATTCAGAGCGAGAACTTAACAAATCACTTAAGAACGAT	588		
DB	481	TTTTCAGCTTAATGGAATTGATTCAGAGCGAGAACTTAACAAATCACTTAAGAACGAT	540		
QY	589	CCGCATATCTTAACTATTAACCGGAATTTTTCAGAGTGAACATACCGGTTCAGTGTG	648		
DB	541	CCGCATATCTTAACTATTAACCGGAATTTTTCAGAGTGAACATACCGGTTCAGTGTG	600		
QY	649	GGAACCGGACTTCGACCGGAAGTGAACAACTGATGCCAATGCGCAACCAACGAGCTTA	708		

Db	Accession	Length	Source
Db	601	660	GGAAACCGGACCTCGGACCGGAAGTGGAAACACAGTGTATGCAACATTCGGCAACCAAGCAACTTA
Qy	709	768	AACCATGAGAGATTTCACAAAGAGTGTAGGCGCAAGAAAGCCATTTCTGTGCTGCGGAT
Db	661	720	AACCATGAGAGATTTCACAAAGAGTGTAGGCGCAAGAAAGCCATTTCTGTGCTGCGGAT
Qy	769	828	TACGGAACATATGATATGTTGGAGCGATGATTTGTCGCCGGTTTGTGTGGGTTTATGGCCGGT
Db	721	780	TACGGAACATATGATATGTTGGAGCGATGATTTGTCGCCGGTTTGTGTGGGTTTATGGCCGGT
Qy	829	888	TGATATGTGTAGAAATGGGCAAGAAAAAAGTCTGAGATGAGAGCTTTGTAGGTGGAATT
Db	781	840	TGATATGTGTAGAAATGGGCAAGAAAAAAGTCTGAGATGAGAGCTTTGTAGGTGGAATT
Qy	889	948	GTGTTTCCGTTTCTCAAGTATAGTTTCTGGGGTGAAGAAAGCGAGATTCATGATGATTTG
Db	841	900	GTGTTTCCGTTTCTCAAGTATAGTTTCTGGGGTGAAGAAAGCGAGATTCATGATGATTTG
Qy	949	1008	AAGATATCCTTCCTGTTCTCCGGCCAGCTGATCTTCACTGATGTTGGAAGAAGCTTCT
Db	901	960	AAGATATCCTTCCTGTTCTCCGGCCAGCTGATCTTCACTGATGTTGGAAGAAGCTTCT
Qy	1009	1023	GGTATCTTCTGCTTNG 1023
Db	961	975	GGTATCTTCTGCTTNG 975
RESULT 11			
LOCUS	AF337544	1115 bp	mRNA linear PLN 22-OCT-2002
DEFINITION	Brassica oleracea chlorophyllase 1 mRNA, complete cds.		
ACCESSION	AF337544		
VERSION	AF337544.1	GI:24210532	
KEYWORDS			
SOURCE			
ORGANISM	Brassica oleracea		
	Brassica oleracea		
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;		
	rosids; eurosids II; Brassicales; Brassicaceae; Brassica.		
REFERENCE	1 (bases 1 to 1115)		
AUTHORS	Chen,H.-H., Chou,Y.-F. and Shaw,J.-F.		
TITLE	Molecular characterization and differential regulation of three		
	chlorophyllase genes, members of the novel serine esterase gene		
	family, in Broccoli		
	Unpublished		
JOURNAL	2 (bases 1 to 1115)		
REFERENCE	Chen,H.-H., Chou,Y.-F. and Shaw,J.-F.		
AUTHORS	Direct Submission		
TITLE	Submitted (18-JUN-2001) Institute of Botany, Academia Sinica,		
JOURNAL	Taipei, Taiwan 11529, Republic of China		
FEATURES	Location/Qualifiers		
source	1..1115		
	/organism="Brassica oleracea"		
	/mol_type="mRNA"		
	/culivar="Green King"		
	/db_xref="taxon:3712"		
	/cfeature_type="feature"		
	1..981		
	/EC_number="3.1.1.14"		
	/note="Boc1H1"		
	/codon_start=1		
	/product="chlorophyllase 1"		
	/protein_id="AA51933.1"		
	/db_xref="GI:24210533"		
	/translation="MAGKESETFSSAATPELAGSLPTTVIPADPSATDLTAPKP		
	VLDSPVAGTVPVLFPHFGVIRNFYSVIVHVAISHGIYVAPOLCKILPQGGVE		
	VIIDGKIVNTSKNLKAHLPSVANGNVTALVSHRGKGKTPAVALLGHAATLDPISX		
	PSALVIGDIPVAGISKIRNDPELITVPPSEFDIDMPVAVIGTGLGPKSNMLAPCPA		
	EVNHEFYIECKATKGFVPAVDYGHMDMDNDIPGVFGVAGCMCKNGKRSKMSF		
	VGGIVAAFLKTSIGWMSERITQLIKDPSVPAFLDPSPELEASGLV"		
	982..1115		

Query Match 63.5%; Score 754; DB 15; Length 1115;
 Beec Local Similarity 82.7%; Pred. No. 4.9e-201;
 Matches 862; Conservative 0; Mismatches 180; Indels 0; Gaps 0;

QY 43 ATACAAATGGCGGAGATAGAGACAGTCACAGTCTTCTCTGTGTAACTCCGGCGGCT 102
 |||||
 DB 1 ATACAAATGGCGGAGAGAGAGACAGTCAGACGTTTCTCGCGGCACTCTTTGGCG 60
 103 TTGAGATAGGAGACCTCCGACACACCGAGATACCGGTGATCCGGTGAATAATGATTC 162
 61 TTGAGTTAGGAGACCTTCCAAACACCGTATCCCGGAGACCGGTGGGACAGGATTTG 120
 QY 163 ACAGACCGGCGGAAACCGGTGAGATATCCTGTCCACAGTCGCCGGAATTTATCCGTC 222
 121 ACCGACCTCCAAAGCCGTATATATACCTCCCAACCGTCCGCGGAATCTACCCGTC 180
 223 GTTTATTTCTTCAGTGGCTTTATCTTGCAACTACTCTACTGAGCTTCTTAACAC 282
 181 GTCTTATTTCTTCAGTGGCTTTATCTTGCAACTACTCTACTGAGCTTCTTAACAC 240
 QY 283 ATCCCTTCGATGCTTACATTTCTGTAGCCCGACAGTTGCAATTAATTCGCCGCG 342
 241 GTACCTTCTAGGCTACATTTGTTAGCCCGACAGCTTTGCAAGATTTTGCGCGGGA 300
 QY 343 GGGCAATGGAAATGAGACAGTCTGGAAGTGTATTAATCTGGGCAATGGAAAACCTCAA 402
 301 GGGCAATGGAAATGAGACAGTCTGGAAGTGTATTAATCTGGGCAATGGAAAACCTCAA 360
 QY 403 GCTACCTACCAACTGGTAAATGCTAATGAAATATACCTCACTGTTGGGCGACAGC 462
 361 GCTACCTACCAACTGGTAAATGCTAATGAAATATACCTCACTGTTGGGCGACAGC 420
 QY 463 CGCGGTGGAAAACCGCGTTTCCGTTGCGGTAGCCCAATGCGCAATTAAGACCAATCC 522
 421 CGCGGTGGAAAACCGCGTTTCCGTTGCGGTAGCCCAATGCGCAATTAAGACCAATCC 480
 QY 523 ATCACTTTTACGCTTATAGAAATGATCCATCGCAGGAATTAACAATATCATTAAGA 582
 481 ATCAAGTTTTCAGCTTGTAGAAATGATCCATCGCAGGAATTAACAATATCATTAAGA 540
 QY 583 ACCGATCCGATATCTTAAAGTATTAACCGGAATCTTTGAGCTGAGCATATCCGGTTGCA 642
 541 ACCGATCCGATATCTTAAAGTATTAACCGGAATCTTTGAGCTGAGCATATCCGGTTGCA 600
 QY 643 GTGTGGGAAACCGGACTCGGACCGAAGTGAACAAGTATGCCATGCGACCAACG 702
 601 GTGTGGGAAACCGGACTCGGACCGAAGTGAACAAGTATGCCATGCGACCAACG 660
 QY 703 GACTTAAACATGAGAGTTTAAAGAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 762
 661 GAACTGAACATGAGAGTTTAAAGAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 720
 QY 763 GCGGATTAACGAGATATGATATGAGAGATGATTTGCGGTTTGTGGTTATG 822
 721 GCGGATTAACGAGATATGATATGAGAGATGATTTGCGGTTTGTGGTTATG 780
 QY 823 GCGGTTGTATGTATGAAGATGAGCAAGAAAAGTCTGAGATGAGAGCTTTGTAGT 882
 781 GCGGTTGTATGTATGAAGATGAGCAAGAAAAGTCTGAGATGAGAGCTTTGTAGT 840
 QY 883 GGAATTTGTGTGCTTTCTCAAGTATATGTTGGGGTGAAAAAGCGAGATTCGATG 942
 841 GGAATTTGTGTGCTTTCTCAAGTATATGTTGGGGTGAAAAAGTCTGATTCGATG 900
 QY 943 ATTGTAAGATCTTCCGTTTCTCGGCGCAAGCTTATCACTCACTGAGTTGGAAGA 1002
 901 ATTGTAAGATCTTCTGTTTCTCGAGCGAGCTTATCTTGGCGGAGCTGGAAGAG 960
 QY 1003 GCTTCTGATCTTCTGCTAGATTTGTATATGATCTATTAATGAGAGGGGCTTGAATAT 1062
 961 GCTTCTGATCTTCTGCTAGATTTGTATATGATCTATTAATGAGAGGGGCTTGAATAT 1020

QY 1063 TTGAAAACCTATCATGTTT 1084
 |||||
 DB 1021 TCCAAACTCGTATTAATCTAT 1042

RESULT 12
 EX510604 453 bp DNA linear STS 10-JUN-2003
 LOCUS Arabidopsis thaliana transposon insertion STS SM_3.40564, sequence tagged site.
 DEFINITION Arabidopsis thaliana (thale cress)
 ACCESSION BX510604
 VERSION BX510604.1 GI:30577550
 KEYWORDS STS; STS, sequence tagged site.
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE
 1. Clarke, J.H., Bowles, B., Carter, J., Hart, D., McCullagh, B., Murphy, G., Langham, S., Legry, C., Jones, J.D.G. and Bevan, M. Unpublished
 2. (bases 1 to 453)
 TITLE Direct Submission
 JOURNAL Submitted (12-MAY-2003) Clarke J.H., John Innes Centre, Colney Lane, Norwich, NR4 7UJ, UK
 COMMENT AT denotes an activation tag disassociation transposon within a single line. ET an enhancer trap disassociation transposon, GT a gene trap disassociation transposon, SM a defective suppressor mutator transposon, _3 denotes a sequence derived from the 3' end of the transposon, _5 denotes a sequence derived from the 5' end of the transposon BBSRC GARNET, ATIS project
 On-line seed stock requests: <http://nasc.nott.ac.uk/> NASC stock code: N127275.

FEATURES
 source location/Qualifiers
 1..453
 /organism="Arabidopsis thaliana"
 /mol_type="genomic DNA"
 /variety="Columbia-0 NASC stock code N1092"
 /db_xref="taxon:3702"
 /clone="AC024609"
 /note="Derived from superpool 19.36 NASC code M40883"
 1..453
 /standard_name="SM_3.40564"

ORIGIN
 Query Match 33.3%; Score 395.8; DB 10; Length 453;
 Best Local Similarity 98.3%; Pred. No. 5.2e-100;
 Matches 400; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 303 TCTGTAGCCCAAGTGTGCAATTAATGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAG 362
 |||||
 DB 47 TATTTCTGAATGACAGTTGTGCAATTAATGCGCGGAGGAGGAGGAGGAGGAGGAGGAG 106
 QY 363 TGCTGAAGTGTATTAATCTGGGATCGGAAAACCTCAAGTCTACCTCAACTTCGCT 422
 107 TGCTGAAGTGTATTAATCTGGGATCGGAAAACCTCAAGTCTACCTCAACTTCGCT 166
 QY 423 AATGTCTAATGAAAATATACCTCACTCGTGGGCGACAGCGGGGTGGAAAACCGCGCTT 482
 167 AATGTCTAATGAAAATATACCTCACTCGTGGGCGACAGCGGGGTGGAAAACCGCGCTT 226
 QY 483 TGGGTTGCGCTAGGCGATGCCCAACATTAAGACCATCATGACGTTTCACTCAAT 542
 227 TGGGTTGCGCTAGGCGATGCCCAACATTAAGACCATCATGACGTTTCACTCAAT 286
 DB 543 AGGAATTTGATCGATCGGAGAACTTAACAATATTAAGAACCGATCCGATATCTTAAC 602
 287 AGGAATTTGATCGATCGGAGAACTTAACAATATTAAGAACCGATCCGATATCTTAAC 346
 QY 603 GTATTAACCGGATCTTTGAGCTGAGCATACCGGTTGCAAGTGTGGAAACCGGACTCG 662

Db 347 GATATACCGGATCTTTGAGCTGAGATACCGTTGCAGTGGGAGACCGAGCTGG 406
Qy 663 ACCGAAGTGAACAACGTGATGCCACCATGGCCACCAACGACTTAA 709
Db 407 ACCGAAGTGAACAACGTGATGCCACCATGGCCACCAACGACTTAA 453

RESULT 13
AX411601
LOCUS AX411601 987 bp DNA linear PAT 14-JUN-2002
DEFINITION Sequence 1 from Patent WO0229022.
ACCESSION AX411601
VERSION AX411601.1 GI:21444159
KEYWORDS
SOURCE
ORGANISM Vitis sp.
Eukaryota; Vitridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; Vitaceae; Vitis.
REFERENCE 1
AUTHORS Thorpe, C., Cahoon, E.B. and Cahoon, R.E.
TITLE Chlorophyllases
JOURNAL Patent: WO 0229022-A 1 11-APR-2002;
E.I. DU PONT DE NEMOURS AND COMPANY (US)
FEATURES
SOURCE 1..987
/organism="Vitis sp."
/mol_type="unassigned DNA"
/db_xref="taxon:3604"

ORIGIN
Query Match 20.2%; Score 239.4; DB 6; Length 987;
Best Local Similarity 57.6%; Pred. No. 6.7e-56;
Matches 473; Conservative 0; Mismatches 336; Indels 12; Gaps 2;
Qy 167 CACCCGCAAAACCGGTGAAGATCACCTGTCCAAACAGTGGCGGAATTATCCGCTTT 226
Db 26 CCCGTCACAGCCATTGTGATGTATGATACCAACCATTCAGGAGCATACCAAGTTCTCT 85
Qy 227 TATTTTCATGCGCTTTTATCTTGGCAACTACTTCTACTGAGCTTTTAAACACATCG 286
Db 86 TGTTCCTTCATGCGCTTTCAGCTCCGCAACACCTTCTACACTGAGCTCTTCAACTCATTT 145
Qy 287 CTGCGATGTTATCATCTTTGTGAGCCCAAGTTGCAAAATTATTTGCGCGCGGAGGCG 346
Db 146 CTTCCCATGATTCATTGTGTGCTCTCTCACTTATACGACTATTTACTCTTCTGAA 205
Qy 347 AAGTGAAGTGAACGATGCTGGAAGTGTATTAACCTGGGCAATCGGAAACCTCAAGCTC 406
Db 206 TTCAAGAGATCAAAATCAGAGCAGATCAAAATTTGGCTATTCCTCAGGCTTCAATCTG 265
Qy 407 AACTTCAACTTCGGTAAATGCTATGGAATTAACCTCACTCGTGGGCCACAGCCGCG 466
Db 266 TGCTCCCAAGAAATGTGAACCAAGACTTCACTCAACCTTCTTTCAGGCCACAGCAGAG 325
Qy 467 GTGGGAAAACGGCGTTTGCGGTTGCGCTAGAGCCACGCAACATTAGAACCCATCATCA 526
Db 326 GGGGAAACAGACATTTGCTGCGCATAGGATATGC-----TATATCATCTCTCA 376
Qy 527 CGTTTCAGCTCTAATAGAAATGATTCAGTGCAGAACTTAAACAAATATATAGAACCG 586
Db 377 ACTTCTCAGCCCTTACTAGACTAGACCTGTGTGGGTGATGATTAATGTGGCCAAACAG 436
Qy 587 ATCCGATATCTTAACGTAATTAACCGGAATCTTTGAGCTGGAATACCGGTGGAGTGG 646
Db 437 TTCCCAAAATCTTAATCTATGTTCTCATTTCTTCAATCTAGCAATCCAGTTGCGTAA 496
Qy 647 TGGGAACGGAGCTGGACCGAAG---TGAACAAAGTATGCGCATGCGACGACCAACG 703
Db 497 TCGGACGGGGGTGGGCGATAGGCGCAAGAACTGCTTAACATGTCTCATGTGCCACAGATG 556
Qy 704 ACTTAACCATGAGAGTTTATCAAAAGATGTAAAGGCGACGAAGCCCATTTCTGTGCTG 763

Db 557 GAGTGAACCATGTAGATTTTTCAGTGAAGTAAACCTCCTTGTTCACCTTTGTGACTA 616
Qy 764 CGATTTCAGCATATGATATGTTGAGCATGATTTGCCGTTTGTGGGTTATGC 823
Db 617 CTGAATATGTGTCACTTGGACATGTATGATATCATCTTTCAAGGTGATGGGCGGATTT 676
Qy 824 CCGGTGTATGTGTAGAAATGGGCAAGAAAAGTCTGATGAGAGAGCTTTAGAGTG 883
Db 677 CGGTTATATCTGCAAGAGTGGAGAGGCTCTAGGACCCCATGAGAGATGTGGGTG 736
Qy 884 GAATTTGTTGCGCTTTCTCAAGTATATGTTGGGGTGAAGAAACCGAGATTCGATTGA 943
Db 737 GCGTTTGTGTTGATCTTCAAGGCTTATTTGGAAGGTCAAGATCGAGATTTCAAAAGCA 796
Qy 944 TTGTGAAGATCCTTCGCTTTCCGCGCAAGCTTATCTCT 984
Db 797 TTGTGTATGAACCTGATGCTGCTCTGTGAAGCTTGTATCTCT 837

RESULT 14
AX411607
LOCUS AX411607 1174 bp DNA linear PAT 14-JUN-2002
DEFINITION Sequence 7 from Patent WO0229022.
ACCESSION AX411607
VERSION AX411607.1 GI:21444162
KEYWORDS
SOURCE Glycine max (soybean)
ORGANISM Glycine max
Eukaryota; Vitridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
REFERENCE 1
AUTHORS Thorpe, C., Cahoon, E.B. and Cahoon, R.E.
TITLE Chlorophyllases
JOURNAL Patent: WO 0229022-A 7 11-APR-2002;
E.I. DU PONT DE NEMOURS AND COMPANY (US)
FEATURES
SOURCE 1..1174
/organism="Glycine max"
/mol_type="unassigned DNA"
/db_xref="taxon:3847"

ORIGIN
Query Match 17.1%; Score 202.6; DB 6; Length 1174;
Best Local Similarity 56.9%; Pred. No. 1.6e-45;
Matches 480; Conservative 0; Mismatches 339; Indels 24; Gaps 5;
Qy 154 AATGATTCACAGACCGCCAAAACCGGTGAGATCACTGTCCAAACAGTCGCCGAACT 213
Db 102 AATGCTTCCTCCTCAGCTCCAAAACATTTGTATCTTTACACCAACCGTGCTGCTCA 161
Qy 214 TATCCGCTGTTTATTTCTTCATGCTTTTATCTTTCGCAACTATCTTCTAGCTT 273
Db 162 TACCTGTATATATGTTCTGCAATGATTTTCTTTCGCAATGACTACTCTGAGCTC 221
Qy 274 CTTAACCAATCGCTTGGCATGTATCTTGTAGCCCAAGTTGRCGAATTTATG 333
Db 222 CTAGGCCAATAGTTTACATGATGATCATATATGTTGCTCTTCAAGCTGTGGAGTGA 281
Qy 334 CCGCC-----GGAGGGCAAGTGAAGTGAAGATCTGGAAGTGAATTAAGTGG--- 384
Db 282 CGGTATATGTTGAACCTGATGATGAAGTTAAATTTGCAGGAAAGTTGTGATTTGCTA 341
Qy 385 GCATGGAAAACCTCAAGCTCACTTCAACAACTTGGTAAATGCTAAATGAAAATACACC 444
Db 342 GCCAGAGAGGGGCTTCAACTCTGCTTCCAGAGATGTGAAGCAATTTGATTAATTTG 401
Qy 445 TCACTGCTGGGCCACAGCCGCGGTGGGAAAACGGCGTTTCGTTGCGTGAAGCCATGCC 504
Db 402 GTTTATACAGTCAACAGAGAGGTGGCAAAACGTATTTGCTGTGCGACTTGTATGCT 461

Qy 505 GCAACATTAGACCATCCATCGTTTACGCTTAATAGAAATTGATCCAGTGGCAGA 564
Db 462 AAAA-----CTAACCTGAAGTTTCAGACCTAGAGCATAGACCTGTGGCTGGC 512
Qy 565 ACTAACAAATACATTAGAACCCGATCCGATCTTAAAGTAAACCGGAATCTTTCGAG 624
Db 513 CCATGTAAATCTTGGGAAACATTTCTCTATTTCTACTGCGCATGTCCCATCTTCAAT 572
Qy 625 CTGACATACCGGTTGCACTGTGGGAAACCGGACTCGGACC--GAAGTGAACACATG 681
Db 573 TTGAACATACCCATTTGTTGATTTGGCACTGGCTAGGCCCAAGAAAGGCTAATTTTTT 632
Qy 682 ATGCACCATGCGCAGCAAGGACTTAAACATAGAGGTTTCAAAAGGTGAAGGCG 741
Db 633 ATTACACATGTGCTCTGATGGGATGAAACATAGAGGTTTCAATAGTGAACACC 692
Qy 742 ACGAAGCCCATTTCTGCTGCTCGGATTAACGACATATGATATGTTGACATGATTTG 801
Db 693 CTTGTGCAATTTTGTGTCAGATGATGTGCAATGACATGTTGATGATGATGACA 752
Qy 802 CCCGCTTTTGTGGGT--TTATGCGCGGTTGATGTTGAAGAAATGGGCAAGAAAAAG 858
Db 753 CTTGGCTTAATTTGGGTCAATTTGTCAATTTGATGCAAGATGGGAAAGGCTCTAGG 812
Qy 859 TCTGATGAGGAGCTTTGTAGTGAATGTGGTGGTTCTCAAGTATAGTTGTGG 918
Db 813 GACTTGAATGAAAGACCGGTGGAGGTTGTTGTGCTTTAAAGGCGACAGTTAAT 872
Qy 919 GGTGAAAAAGCGAGATTCGATTTGATTTGTAAGAGATCTTCCGTTCCGCGCAAGCTT 978
Db 873 GGCTATGGAAGATTTTAATGCTGTTTGGCGAATCTTAATCTTGTCTCTCACTAACTG 932
Qy 979 GAT 981
Db 933 GAT 935

RESULT 15
AX411617 1242 bp DNA linear PAT 14-JUN-2002
LOCUS Sequence 17 from Patent WO229022.
DEFINITION AX411617
ACCESSION AX411617
VERSION AX411617.1 GI:21444167
KEYWORDS
SOURCE Triticum aestivum (bread wheat)
ORGANISM Triticum aestivum
REFERENCE 1
AUTHORS Thorpe, C., Cahoon, E.B. and Cahoon, R.E.
TITLE Chlorophyllases
JOURNAL Patent: WO 0229022-A 17 11-APR-2002;
E.I. DU PONT DE NEMOURS AND COMPANY (US)
FEATURES
location/Qualifiers
source 1..1242
/organism="Triticum aestivum"
/mol_type="unassigned DNA"
/db_xref="taxon:4565"

ORIGIN
Query Match 16.8%; Score 199; DB 6; Length 1242;
Best Local Similarity 55.8%; Pred. No. 1,7e-44;
Matches 476; Conservative 0; Mismatches 350; Indels 27; Gaps 4;

Qy 137 CGGTGATCCGGTGAATAATGATTCACACAGACCGCAAAACCGGTGAGATCACTGTC 196
Db 153 CGATTCAGTGTGATGAAATGCGGCGCGACACACCGATCCGGTCTGATGTCGAC 212
Qy 197 CAACAGTCGCGGAATTAATCCGCTGTTTATTTTCATGAGCTTTATCTTGCAACT 256
Db 213 CCAAGATGAGGAACCTACCCCGTGGCAATGCTTTCAGAGGCTTCTCTCATTAAC 272

Qy 257 ACTTTACTGAGCTTTTAAACCAATGCGCTTTCGATGTTACATTTCTGTAGCCCCAC 316
Db 273 ACTTCTAGCAACCTTCTCCGACAGTGCATTCACACGCTTTCATCATTTGTGCGCCCC 332
Qy 317 AGTTGTTG---CAATATATTCGCCCGGAGGAGGCAAGTGAAGTGAAGTGAAGT 373
Db 333 AGTTGACATCAATATATATCTTCCGGTGAACGAGACATCGCCGCGGACCAAG 392
Qy 374 TGAATACTGGGATCGGAATAACCTCAAGCTCACTTCAACTTGGTAAATGTAATG 433
Db 393 TGGCAAGATGCTGCCCGAGCGGCTCCGTCGTCGCCAAGAGGATGAGCCGAGC 452
Qy 434 GAAATATACCTCACTGTGGGCAACAGCCGCTGGGAAAGAGGCTTGGCTGGCC 493
Db 453 TCTGAAGCTGCTTGGCGGCAACAGCGAGAGGCCACAGGCTTCTCCCTGGCT 512
Qy 494 TAGGCAATGCGGCAACTTGAACCATTCATCACTGTTTCACTTAATGAATTGATC 553
Db 513 TGGGGACGCGC-----AAGATCCAGTAACTTCTCCGCTCATGAGCTGAC 563
Qy 554 CAGTGCAGGAACCTAACAAATACATTAGAAACGATCCGATATCTTAAAGTAAACCG 613
Db 564 CCGTCCCGGCAAGGGAAGTCTCCAGGCTCCAGCCCAAGATCTCACTTCAAGGCGT 623
Qy 614 AATCTTTCAGCTGACATACCGGTTGCACTGATGGAACCGGACTTCGACCGAAGTGA 673
Db 624 CTTCTTCGCAATGGGAGATGCTCGTGTGATTCGCAACCGGCTCGGGAAGAAAG 683
Qy 674 ACAACGATG---CCACCATGCGCAACACGAGCTTAAACATGAGAGTTTAAAG 730
Db 684 AGAATATATTTCTTCTCTCCCTGCGCACCAAGACGTGAACACGCGAGTTTACCGCG 743
Qy 731 AGTGAAGGCAAGAAAGCCATTTCTGCTGCGGATTAACGACATATGATATGTTGG 790
Db 744 AGTGAAGGCGGCTGCTACTACTTTGTGACCAAGACATACGCGCATGTGACATGCTGG 803
Qy 791 ACGATATTTGCCCCGTTTGTGGTTTATGAGCGGCTGTATGTTGAAGATGGGCA 850
Db 804 ACGACGACGCCCCCAAGTTTAT-----CACCTGCGTCTCAAGATGGGAAAGC 851
Qy 851 GAAAAAGTCTGATGAGAGCTTTGATGATGGAATTTGCTTCTCAAGTATA 910
Db 852 GGTGCAAGGCAAGATGCGAGAGTGTCTGGAATCATGTGGCATTTCTTAATGCTG 911
Qy 911 GTTTGGGGTGAAGAAAGCGAGATTGATTAATGTAAGAGATCTTCGTTCTCCGG 970
Db 912 CTTGGGTTGAAAGATGAGATCTTGAAGCCATCTGAGAGACCGCGGCTTGACCCA 971
Qy 971 CCAAGCTGATCC 983
Db 972 CCAAGCTGATCC 984

Search completed: March 20, 2006, 16:11:51
Job time : 4208.73 secs

THIS PAGE LEFT BLANK

/db_xref="taxon:3702"
/clone="GSLT8B532C07"
/feature_type="flowers and buds"
/ecotype="COL-0"
/plasmid="pCMVSPORT_6"
complement(1..1169)
/gene="Atlg19670"

ORIGIN

Query Match 71.5%; Score 849.8; DB 4; Length 1169;
Best Local Similarity 86.7%; Pred. No. 4,8e-238;
Matches 1014; Conservative 0; Mismatches 2; Indels 153; Gaps 1;

```

Oy 25 CTTTAAATATCATGTTTATATACAAATGGCGGAGATAGAGACAGTCCAAAGTTTCTCT 84
Db 1 CTTTAAATATCATGTTTATATACAAATGGCGGAGATAGAGACAGTCCAAAGTTTCTCT 60
Oy 85 GTGTAACCTCCGCGGCTTTTGAATAGGACAGCTCCGACAAACGAGATACCGGTGAT 144
Db 61 GTGTAACCTCCGCGGCTTTTGAATAGGACAGCTCCGACAAACGAGATACCGGTGAT 120
Oy 145 CCGGTGAAAAATGATTTCAACAGACACCGCCAAAACCGGTGAGATACCTGTCCAAAGT 204
Db 121 CCGGTGAAAAATGATTTCAACAGACACCGCCAAAACCGGTGAGATACCTGTCCAAAGT 180
Oy 205 GCCGGAACCTTACCCGCTGTTTATCTTCCATGGCTTTTATCTTCCGAACTACTTCTAC 264
Db 181 GCCGGAACCTTACCCGCTGTTTATCTTCCATGGCTTTTATCTTCCGAACTACTTCTAC 240
Oy 265 TCTGACGTTTCTTAAACCAATCGCTTCGATGGTTTACATTTCTTGTAGCCCC----- 314
Db 241 TCTGACGTTTCTTAAACCAATCGCTTCGATGGTTTACATTTCTTGTAGCCCCAAGTACAT 300
Oy 315 ----- 314
Db 301 AATTTACTATACAAAAATAATATATGATCATTTTAAAGATTAACCTTACGATATGTAAC 360
Oy 315 ----- 314
Db 361 ATGTAAATATGTAACAAAAATGATTAATGATTAATGATTAATGATTAATGATTAATG 420
Oy 315 -----ACAGTTGTGCAAAATTAATTTGCGCGCGGAGAGGCAAGTG 351
Db 421 AGTAATATTTTATATTTCTTGAATGATGATGTCATTAATTTATTTGCGCGGAGAGGCAAGTG 480
Oy 352 GAAAGTGAACGATGCTGGAAGTGTGATTAATCTGGGATTCGGAACCACTCAAGCTCACTTA 411
Db 481 GAAAGTGAACGATGCTGGAAGTGTGATTAATCTGGGATTCGGAACCACTCAAGCTCACTTA 540
Oy 412 CCAACTTGGGTAAATGCTTAATGGAATAATCACTCACTCGTGGGCGACAGCGCGGTGGG 471
Db 541 CCAACTTGGGTAAATGCTTAATGGAATAATCACTCACTCGTGGGCGACAGCGCGGTGGG 600
Oy 472 AAAACGGCGTTTGGCGGTGGCGGTAGGCGCATGCGCAACATTAAGACCAATCCATCGTTT 531
Db 601 AAAACGGCGTTTGGCGGTGGCGGTAGGCGCATGCGCAACATTAAGACCAATCCATCGTTT 660
Oy 532 TCAGCTCTTAATGGAATGATTCAGTCCGAGAACTTAACAAATATATTAATGAACGATCCG 591
Db 661 TCAGCTCTTAATGGAATGATTCAGTCCGAGAACTTAACAAATATATTAATGAACGATCCG 720
Oy 592 CATATCTTAACGTAATAACCGGAATCTTTCAGCTGACATATCCGTTGCGAGTGTGGGA 651
Db 721 CATATCTTAACGTAATAACCGGAATCTTTCAGCTGACATATCCGTTGCGAGTGTGGGA 780
Oy 652 ACCGGAAGTGGACCGAAGTGAACCAAGTATGCAACATGCGCAACGAGACTTAAAC 711
Db 781 ACCGGAAGTGGACCGAAGTGAACCAAGTATGCAACATGCGCAACGAGACTTAAAC 840
Oy 712 CATGAGAGTTTAAAGAGATGTAAAGGAGAGAAAGCCATTTGTGGCTGCGGATTAAC 771
Db 841 CATGAGAGTTTAAAGAGATGTAAAGGAGAGAAAGCCATTTGTGGCTGCGGATTAAC 900

```

```

Oy 772 GGACATATGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 831
Db 901 GGAATATGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
Oy 832 ATGTGTAAGAAATGGGCAAAAGAAAAAGTCTGAGATGAGAGCTTTGATGATGATGATGATG 891
Db 961 ATGTGTAAGAAATGGGCAAAAGAAAAAGTCTGAGATGAGAGCTTTGATGATGATGATGATG 1020
Oy 892 GTTGGCTTTCTCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 951
Db 1021 GTTGGCTTTCTCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1080
Oy 952 GATCCTTCCGTTTCTCCGCGCAACCTGATCTTCACTGAGTTGGAAGAAAGCTTCTGAT 1011
Db 1081 GATCCTTCCGTTTCTCCGCGCAACCTGATCTTCACTGAGTTGGAAGAAAGCTTCTGAT 1140
Oy 1012 ATCTTGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1040
Db 1141 ATCTTGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1169

```

RESULT 2
B0636473 796 bp mRNA linear EST 23-SEP-2002
LOCUS 006A06 Infected Arabidopsis Leaf Arabidopsis thaliana cDNA, mRNA
DEFINITION
B0636473
VERSION
KEYWORDS
SOURCE
ORGANISM
Arabidopsis thaliana (thale cress)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 796)
Lundsgaard,M., Emmersen,J., Nielsen,K.L., Wilson,I., Somerville,S.
and Weidner,K.G.
EST sequencing of Erysiphe cichoracearum infected Arabidopsis
plants

TITLE
JOURNAL
COMMENT
Unpublished (2002)
Contact: Karen G. Weidner
Institut for bioteknologi
Aalborg Universitet
Sohnsgaardsholmsvej 49, 9000 Aalborg, Denmark
Tel: +45 96358467
Fax: +45 98141808
Email: Kgwe@bio.auc.dk.
Location/Qualifiers

FEATURES
source
1..796
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/ecotype="Columbia"
/db_xref="taxon:3702"
/dev_stage="Plant 3 weeks old, three days post infection"
/clone_id="Infected Arabidopsis Leaf"
/note="Organ: Leaf; Vector: plasmid; Mixed cDNA
library of Arabidopsis and E. cichoracearum infected leaf
from three weeks old Arabidopsis plants. Plants were
harvested 3 days after infection and mRNA oligo dt
selected."

ORIGIN

Query Match 65.1%; Score 772.8; DB 5; Length 796;
Best Local Similarity 99.5%; Pred. No. 1.9e-215;
Matches 786; Conservative 0; Mismatches 2; Indels 2; Gaps 1;

```

Oy 30 ATTATCTAGTTTAATACAAATGGCGGATAGAGACAGTCCAAAGTTTCTCTGCTG 89
Db 7 ATTATCTAGTTTAATACAAATGGCGGATAGAGACAGTCCAAAGTTTCTCTGCTG 66
Oy 90 AACTCCGGCGGCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 149
Db 67 AACTCCGGCGGCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 126

```

Qy 150 GGAATATGATTCACAGCAGCCGCAAAACCGGTGAGATCACTGTCCAAAGTCGCCG 209
Db 127 GAAAAATGATTCACAGCAGCCGCAAAACCGGTGAGATCACTGTCCAAAGTCGCCG 186
Qy 210 AACTATCCGCTGTTTATTCATTCATGAGCTTTATTCGCAACTACTCTGCTGCA 269
Db 187 AACTATCCGCTGTTTATTCATTCATGAGCTTTATTCGCAACTACTCTGCTGCA 246
Qy 270 CGTTCTTAACCAATCGCTTTCATGAGTTCATTCCTG--TAGCCCAACAGTTGCAAA 327
Db 247 CGTTCTTAACCAATCGCTTTCATGAGTTCATTCCTGTTAGCCCAACAGTTGCAAA 306
Qy 328 TTATTCGCGCGGAGGAGGAGAGTGAAGTGAACATGCTGAAGTGTGATTAATCGGCA 387
Db 307 TTATTCGCGCGGAGGAGGAGAGTGAAGTGAACATGCTGAAGTGTGATTAATCGGCA 366
Qy 388 TCGGAAACCTCAAGCTCACTCACTCACTCGGTAATGCTAATGGAATAACACTCA 447
Db 367 TCGGAAACCTCAAGCTCACTCACTCACTCGGTAATGCTAATGGAATAACACTCA 426
Qy 448 CTGCTGAGCCACAGCCGCGGTGAGAAACCGCGTTTGCGGCTAGGCCATGCCCA 507
Db 427 CTGCTGAGCCACAGCCGCGGTGAGAAACCGCGTTTGCGGCTAGGCCATGCCCA 486
Qy 508 ACATTAACCCATTCATCACTGTTTCACTTAATGAAATTCATCCAGTCCGCAACT 567
Db 487 ACATTAACCCATTCATCACTGTTTCACTTAATGAAATTCATCCAGTCCGCAACT 546
Qy 568 AACAAATACATTAAGACCGATCCGATATCTTAAGTAAACCGGAATCTTTCAGCTG 627
Db 547 AACAAATACATTAAGACCGATCCGATATCTTAAGTAAACCGGAATCTTTCAGCTG 606
Qy 628 GACATACCGGTGAGTGTGAGAAACCGGATCCGACCGAAGTGAACAAAGTATGCCA 687
Db 607 GACATACCGGTGAGTGTGAGAAACCGGATCCGACCGAAGTGAACAAAGTATGCCA 666
Qy 688 CCAATGCCACCAACGAGACTTAACATGAGAGAGTTTAAACAAAGTGAAGGCAAGAA 747
Db 667 CCAATGCCACCAACGAGACTTAACATGAGAGAGTTTAAACAAAGTGAAGGCAAGAA 726
Qy 748 GCCCATTTGCTGCTGCGGATTAACGACATATGATGATGTTGACGATGTTGCCGCT 807
Db 727 GCCCATTTGCTGCTGCGGATTAACGACATATGATGATGTTGACGATGTTGCCGCT 786
Qy 808 TTGTGTGGGT 817
Db 787 TTGTGTGGGT 796

RESULT 3
CB255908
LOCUS
DEFINITION
12-EO12744-027-007-H04-T7R MPZ-ADIS-027 Arabidopsis thaliana cDNA
CB255908
VERSION
CB255908.1 GI:32880681
SOURCE
Arabidopsis thaliana (thale cress)
ORGANISM
Arabidopsis thaliana
Eukaryotes; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
REFERENCE
1 (bases 1 to 699)
Schmid, K.J., Soerensen, T.R., Stracke, R., Torjek, O., Alemann, T.,
Mitchell-Olds, T. and Weisshaar, B.
Large-scale identification and analysis of genome-wide
single-nucleotide polymorphisms for mapping in Arabidopsis thaliana
Genome Res. 13 (6), 1250-1257 (2003)
12793357
JOURNAL
PUBMED
Contact: Weisshaar B
ADIS DNA core facility at MPZ
Max-Planck-Institute for Plant Breeding Research

Carl-von-Linne Weg 10, 50829 Koeln, Germany
Fax: 00492215062851
Email: weisshaar@mpz-koeln.mpg.de
Insert Length: 699 Std Error: 0.00
Plate: 7 row: H column: 04
Seq primer: T7R; CTAATACGACTCACTATAGGA.
Location/Qualifier

FEATURES
Source

1. 699
/organism="Arabidopsis thaliana"
/mol_type="RNA"
/ecotype="E1-2"
/db_xref="GABI:592398"
/db_xref="taxon:3702"
/clone="MPZp72H0470"
/tissue_type="whole plant"
/dev_stage="adult plant, mixed stresses"
/lab_host="E. coli TOP10"
/clone_11b="MPZ-ADIS-027"

/note="Vector: pSPORT1; Site 1: SalI; Site 2: NotI; cDNA library from Arabidopsis thaliana, accession Eifel-2; ten week old total plants grown under long-day conditions in soil, whole adult plants were treated for 24 hours with different stresses" (1) at 4M-0 C in the dark, (2) at 37 Grad C in the dark, (3) lying in the lab after removing from soil, (4) in the greenhouse after wounding leaves with a forceps, (5) in the lab watering with a 150 mM NaCl solution, (6) at 26 M-0C in the light/UV; equal quantities of stressed plant material were pooled; library was made at the Max-Planck-Institute for Plant Breeding Research, Cologne, Germany; cloning sites SalI-NotI, primer sites and orientation:
T7-SalI-CCACGCGTCC-5prime-cDNA-polyA-CC-NotI-Sp6; Note: Sequencing granted in the context of the GABI Arabidopsis Verbund I: Genetic Diversity, 'Establishment of high-efficiency SNP-based mapping tools and development of methods for genome-wide mutation detection'; PI: Bernd Weisshaar Sequence submission managed by RZPD/GABI-Primary database: <http://gabi.rzpd.de>. This clone is available from RZPD; contact RZPD (clone@rzpd.de) for further information."

ORIGIN

Query Match 58.7%; Score 697.4; DB 6; Length 699;
Best Local Similarity 99.9%; Pred. No. 2; e-193;
Matches 698; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
Qy 18 CACAACCTTTAATTAATCTAGTTTAATACAAATGGCGGATAGAGACAGTCCAAAGT 77
Db 1 CACAACCTTTAATTAATCTAGTTTAATACAAATGGCGGATAGAGACAGTCCAAAGT 60
Qy 78 TTCTCTGTGTAACTCCGCGGCTTTTGAGATGAGCAGCTCCGACAAACGAGATACC 137
Db 61 TTCTCTGTGTAACTCCGCGGCTTTTGAGATGAGCAGCTCCGACAAACGAGATACC 120
Qy 138 GGTGATCCGCTGAGAAATGATTCACACAGCAGCCGCAAAACCGGTGAGATCACTGTC 197
Db 121 GGTGATCCGCTGAGAAATGATTCACACAGCAGCCGCAAAACCGGTGAGATCACTGTC 180
Qy 198 AACAGTCCGCGGAACCTATCCGCTGTTTATTCATTCATGAGTTCATTCCTTGCAAGTA 257
Db 181 AACAGTCCGCGGAACCTATCCGCTGTTTATTCATTCATGAGTTCATTCCTTGCAAGTA 240
Qy 241 CTCTACTGACGTTCTTAACACATGCTTCGATGTTACATTCCTGTAGCCCCACAC 300
Db 258 CTCTACTGACGTTCTTAACACATGCTTCGATGTTACATTCCTGTAGCCCCACAC 317
Qy 318 GTTGTGCAATTAATTCGCGCGGAGGAGGAGAGTGAAGTGAAGTGTGAT 377
Db 301 GTTGTGCAATTAATTCGCGCGGAGGAGGAGAGTGAAGTGAAGTGTGAT 360
Qy 378 AAACCTGGCATCGGAAACCTCAAAAGCTACCTCAACCTTCGCTGAATGTGAAGA 437
Db 361 AAACCTGGCATCGGAAACCTCAAAAGCTACCTCAACCTTCGCTGAATGTGAAGA 420

Qy 438 ATACAGCTCAGTCGTCGAGGCGACAGCGGCGTGGGAAAAAGCGCTTTGCGCTAGG 497
Db 421 ATACAGCTCAGTCGTCGAGGCGACAGCGGCGTGGGAAAAAGCGCTTTGCGCTAGG 480
Qy 498 CCATCCCGCAACATTAGAACCCATCCATCAGCTTTTCACTTAATAGGAATTCAGT 557
Db 481 CCATCCCGCAACATTAGAACCCATCCATCAGCTTTTCACTTAATAGGAATTCAGT 540
Qy 558 CGCAGGAATCAACAAATACATTAGAACCCATCCATCAGCTTTTCACTTAATAGGAATTCAGT 617
Db 541 CGCAGGAATCAACAAATACATTAGAACCCATCCATCAGCTTTTCACTTAATAGGAATTCAGT 600
Qy 618 TTTCAGCTGACATACCGGTTTGCAGTGGGAAACCGGACTCGGACCGAAATGGGAA 677
Db 601 TTTCAGCTGACATACCGGTTTGCAGTGGGAAACCGGACTCGGACCGAAATGGGAA 660
Qy 678 CGTATGCCACCATGCGACCAACCGACTTAAACCATGA 716
Db 661 CGTATGCCACCATGCGACCAACCGACTTAAACCATGA 699

RESULT 4
BE039090/c 715 bp mRNA linear EST 07-JUN-2000
LOCUS AB09E05 AB Arabidopsis thaliana cDNA 5' similar to
DEFINITION Arabidopsis thaliana cDNA 5' similar to
ACCESSION BE039090
VERSION BE039090.1 GI:8334106
KEYWORDS EST,
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana

REFERENCE
AUTHORS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 715)
Bohner, H.J., Borchert, C., Brazille, S., Brooks, J., Eaton, M.,
Ferreira, H., Kawasaki, S., McCollough, A., Michalowski, C.B.,
Palacio, C., Scara, G., Wheeler, M. and Zepeda, G.R.
TITLE Functional Genomics of Plant Stress Tolerance
JOURNAL Unpublished (2000)
COMMENT Contact: Michalowski, C.B.
University of Arizona
Bio Sciences West room 513, Tucson, AZ 85721, USA
Tel: 520-621-7982
Fax: 520-621-1697
Email: cbm@u.arizona.edu.

FEATURES
source
1. 715
Location/Qualifiers
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/ecotype="Columbia"
/db_xref="taxon:3702"
/issue_type="leaves, seedlings"
/dev_stage="2-3 weeks"
/clone_id="AB"
/note="200mM NaCl"

ORIGIN
Query Match 51.0%; Score 605.6; DB 2; Length 715;
Best Local Similarity 97.4%; Pred. No. 2.7e-166;
Matches 647; Conservative 0; Mismatches 14; Indels 3; Gaps 3;

Qy 527 CGTTTCAGCTTAATAGGAATTCAGTCCGAGGAATCAACAAATACATTAGAACCG 586
Db 675 CGTTTCAGCTTAATAGGAATTCAGTCCGAGGAATCAACAAATACATTAGAACCG 617
Qy 587 ATCC-GCATATCTTAACGATTAACCGGA-ATCTTCAGCTGACATACCGGTGCAGT 644
Db 616 TTCGCGCATATCTTAACGATTAACCGGAATCTTCAGCTGACATACCGGTGCAGT 557
Qy 645 GGTGGAAACCGGACTCGAGCCGAAGTGAACCAAGTATGCCATCGGACCAACCGGA 704
|||||

Db 556 GGTGGAAACCGGACTCGGACCGAAGTGGAACAAGTAATGCACCATCGGACCAACCGGA 497
Qy 705 CTTTAACCATGAGGAGATTTCACAAAGGTGTAGGCGACGAAAGCCATTTCGTGCTGC 764
Db 496 CTTTAACCATGAGGAGATTTCACAAAGGTGTAGGCGACGAAAGCCATTTCGTGCTGC 437
Qy 765 GGATTTACGACATATGATATGTTGACGATATTTCCCGGTTTGTGGGTTTATGCG 824
Db 436 GGATTTACGACATATGATATGTTGACGATATTTCCCGGTTTGTGGGTTTATGCG 377
Qy 825 CGTTTATGTATGTAAGATGGGCAAGAAAAAGTCTGAATGAGAGCTTTGTAGGTGG 884
Db 376 CGTTTATGTATGTAAGATGGGCAAGAAAAAGTCTGAATGAGAGCTTTGTAGGTGG 317
Qy 885 AATTGTGTGTCGTTTCTCAAGTATGTTGTGGGGTGAAGGAGATTCATGAT 944
Db 316 AATTGTGTGTCGTTTCTCAAGTATGTTGTGGGGTGAAGGAGATTCATGAT 257
Qy 945 TGTGAAGATCCTTCGCTTCGCGGCAAGCTTGATCTTCACTGAGTTGGAAGAAGC 1004
Db 256 TGTGAAGATCCTTCGCTTCGCGGCAAGCTTGATCTTCACTGAGTTGGAAGAAGC 197
Qy 1005 TTCTGTATCTTCTGTATGATTTGTATGATCTATATCAAGGGGCTTGAATATTT 1064
Db 196 TTCTGTATCTTCTGTATGATTTGTATGATCTATATCAAGGGGCTTGAATATTT 137
Qy 1065 GAAACCATGATATGTTTCTGATCCAGCTGATGTTGATGTTGATGTTGATGTTGAT 1124
Db 136 GAAACCATGATATGTTTCTGATCCAGCTGATGTTGATGTTGATGTTGATGTTGAT 77
Qy 1125 GTGTATTTTATTAATCTGATCAACAAACATTTGTATGTTTACCCCAAAAAA 1184
Db 76 GTGTATTTTATTAATCTGATCAACAAACATTTGTATGTTTATTAATGTTTAAAAA 17
Qy 1185 AAAA 1188
Db 16 AAAA 13

RESULT 5
CF773835/c 563 bp mRNA linear EST 14-APR-2004
LOCUS AG_FSL_23D02 Arabidopsis ag-1 35S:AG-GR forward subreaction library
DEFINITION Arabidopsis thaliana cDNA clone 23D02, mRNA sequence.
ACCESSION CF773835
VERSION CF773835.1 GI:46368801
KEYWORDS EST,
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 563)
Hu, W., Wang, Y., Bowers, C. and Ma, H.
TITLE Isolation, sequence analysis, and expression studies of florally
expressed cDNAs in Arabidopsis
JOURNAL Plant Mol. Biol. 53 (4), 545-563 (2003)
COMMENT Contact: Hong Ma
Hong Ma lab
The Pennsylvania State University
315 Warrick Laboratory, University Park, PA 16802, USA
Tel: 8148636414
Fax: 8148631357
Email: hxm16@psu.edu
Seq primer: M13 Universal
High quality sequence stop: 563
POLYA=No.

FEATURES
source
1. 563
Location/Qualifiers
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/ecotype="Landsberg erecta"
/db_xref="taxon:3702"

/clone="23D02"
/issue_type="inflorescence lacking open or older flowers"
/dev_stage="4 week old plants"
/clone_1lb="Arabidopsis ag-1 35S:AG-GR forward subtraction library"
/note="Vector: The pT-Adv vector in the Advantage Cloning Kit (CLONTECH); Subtractive library constructed using the CLONTECH PCR-select cDNA subtraction kit. Transgenic plants homozygous for the ag-1 mutation and carrying a 35S:AG-GR construct were generated. RNA was isolated from inflorescences with floral meristems and unopened floral buds and without open or older flowers from transgenic plants that were treated with dexamethasone for 3 days, and used for synthesis of the tester cDNA. RNA from the inflorescences of transgenic plants treated with a control solution without dexamethasone was used to synthesize the driver cDNA. It was observed that a 3-day treatment with dexamethasone can induce ather and carpel characteristics in the flowers of the ag-1 35S:AG-GR plants."

ORIGIN

Query Match 46.1%; Score 548; DB 6; Length 563;
Best Local Similarity 99.1%; Pred. No. 2.2e-149;
Matches 551; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```
QY 481 TTTCGGGTGGCGTACGATCCGCGCAATTAACCATCATCAGCTTTTACGCTCTA 540
DB 556 TTGCGGGTGGCGTACGATCCGCGCAATTAACCATCATCAGCTTTTACGCTCTA 497
QY 541 ATAGAAATGATCCAGTCGACGAACTAACAAATATATAGAACCGATCCGATATCTTA 600
DB 496 ATAGAAATGATCCAGTCGACGAACTAACAAATATATAGAACCGATCCGATATCTTA 437
QY 601 ACATATAACCGGAATCTTTGAGCTGACATACCGGTTGAGTGGGAAACCGGATC 660
DB 436 ATGATATAACCGGAATCTTTGAGCTGACATACCGGTTGAGTGGGAAACCGGATC 377
QY 661 GGACCGAAGTGAACCAACGTATATCCCATATCCGACCAACGACCTTAACCATAGAG 720
DB 376 GGACCGAAGTGAACCAACGTATATCCCATATCCGACCAACGACCTTAACCATAGAG 317
QY 721 TTTTCAAAAGAGTGAAGCGAAGAACCATTTGCGTGGCGATTAACCGACATATG 780
DB 316 TTTTCAAAAGAGTGAAGCGAAGAACCATTTGCGTGGCGATTAACCGACATATG 257
QY 781 GATATGTTGACGATATTTGCGGTTTGTGGTTATGCGCGGTTGATGTATG 840
DB 256 GATATGTTGACGATATTTGCGGTTTGTGGTTATGCGCGGTTGATGTATG 197
QY 841 AATGGCGAAGAAAAGTCTGAGATGAGAGCTTTGATGATGATTTGGTTGCGTT 900
DB 196 AATGGCGAAGAAAAGTCTGAGATGAGAGCTTTGATGATGATTTGGTTGCGTT 137
QY 901 CTCAAGATATGTTGTTGGGGTGAAGAAAGCGGAGATTGATGATGAGGATCTTCC 960
DB 136 CTCAAGATATGTTGTTGGGGTGAAGAAAGCGGAGATTGATGATGAGGATCTTCC 77
QY 961 GTTTCTCGGCGCAAGCTTATCTTCACTGAGTTGAGAGAGCTTCTGATCTTGC 1020
DB 76 GTTTCTCGGCGCAAGCTTATCTTCACTGAGTTGAGAGAGCTTCTGATCTTGC 17
QY 1021 TAGATTTGTGTATGT 1036
DB 16 TAGATTTGTGTATGT 1
```

RESULT 6
LOCUS CM836771 681 bp DNA linear GSS 26-NOV-2004
DEFINITION ET8513.Ds.05.20.01.JU94.b.681 Arabidopsis thaliana Landsberg Ds insertion lines Arabidopsis thaliana genomic clone ET8513, genomic survey sequence.
ACCESSION CM836771

VERSION CM836771.1 GI:56092561

KEYWORDS GSS.
SOURCE Arabidopsis thaliana (thale cress)

ORGANISM Arabidopsis thaliana

REFERENCE 1 (bases 1 to 681)
May, B.P., Smorowski, J., Arroyo, J.-M., Vaughn, M.W., Shen, R., McCombie, W.R. and Martienssen, R.A.
Arabidopsis genomic sequences flanking Ds enhancer and gene traps in transgenic lines

AUTHORS

TITLE

JOURNAL

COMMENT

Unpublished (2004)
Contact: Martienssen RA
Cold Spring Harbor Laboratory
1 Bungtown Rd., Cold Spring Harbor, NY 11724, USA
Tel: 516 367 8322
Fax: 516 367 8369
Email: martiens@cshl.org
This sequence flanks a Ds transposon carrying an enhancer trap in line ET8513. The transposon is located within At1g19670.
Class: transposon-tagged.
Location/Qualifiers

FEATURES

source

1..681

/organism="Arabidopsis thaliana"

/mol_type="genomic DNA"

/ecotype="Landsberg erecta"

/db_xref="taxon:3702"

/clone_1lb="Arabidopsis thaliana Landsberg Ds insertion lines"

/note="Lines of Arabidopsis thaliana were generated which each contain a Ds transposon carrying a glucuronidase reporter gene. Genomic DNA flanking the transposon insertion in each line was amplified by PCR and directly sequenced. More information is available at <http://genetrapp.cshl.edu>."

ORIGIN

Query Match 44.9%; Score 533.8; DB 10; Length 681;
Best Local Similarity 98.7%; Pred. No. 3.5e-145;
Matches 538; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```
QY 303 TCTTGAACCCCAAGTTGTGCAATTAATTCGCGGAGGCGCAAGTGAAGTGAAGA 362
DB 137 TATTTCTGATGACAGTTGTGCAATTAATTCGCGGAGGCGCAAGTGAAGTGAAGA 196
QY 363 TGCCTGAGAGTATTAACCTGCGGCAATCGGAAACCTCAACCTCACTCACTCGGT 422
DB 197 TGCCTGAGAGTATTAACCTGCGGCAATCGGAAACCTCAACCTCACTCACTCGGT 256
QY 423 AATATGCTAATGAAATATACCTCATCTGCGGCGCAACGCGGCGTGGGAAAACGGCGTT 482
DB 257 AATATGCTAATGAAATATACCTCATCTGCGGCGCAACGCGGCGTGGGAAAACGGCGTT 316
QY 483 TCGGTTGGCTGAGCCATGCGCAACATTTAGACCATTCATGACGTTTACCTTAAT 542
DB 317 TCGGTTGGCTGAGCCATGCGCAACATTTAGACCATTCATGACGTTTACCTTAAT 376
QY 543 AGGAATGATCCAGTGGCGAAGCTTAACAAATCATTTAGAACCGATCCGATATCTTAAC 602
DB 377 AGGAATGATCCAGTGGCGAAGCTTAACAAATCATTTAGAACCGATCCGATATCTTAAC 436
QY 603 GTATTAACCGGAATCTTTGAGCTGACATACCGGTTGAGTGTGGAAACCGACTCGG 662
DB 437 GTATTAACCGGAATCTTTGAGCTGACATACCGGTTGAGTGTGGAAACCGACTCGG 496
QY 663 ACCGAAGTGAACAGTGTATGCAATGCGGCAACGACGAACTTAACCATAGAGAGTT 722
DB 497 ACCGAAGTGAACAGTGTATGCAATGCGGCAACGACGAACTTAACCATAGAGAGTT 556
QY 723 TTCAAAAGAGTGAAGCGGCAAGCAATTTGCGGCTGGCGGATTAAGGACATATGGA 782
```

Db 557 TTACAAAGATGTAAAGCGACGAAGCCATTTCGTGGGTGGATTCAGACATATGA 616

Qy 783 TATCTTGAACCATGATTTGCCCCGTTTGTGGTTATGGCCGTTGTATGTATGA 842

Db 617 TATGTTGACGATGATTTGCCGCTTTTGTGGTTATGGCCGTTGTATGTATGA 676

Qy 843 TGGGC 847

Db 677 TGGGC 681

RESULT 7
AV525625 525 bp mRNA linear EST 18-PEB-2004
LOCUS AV525625 Arabidopsis thaliana aboveground organs two to six-week
DEFINITION Old Arabidopsis thaliana cDNA APD27d04r 5', mRNA sequence.
ACCESSION AV525625
VERSION AV525625.1 GI:8685153
KEYWORDS EST.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 525)
Asamizu, E., Nakamura, Y., Sato, S. and Tabata, S.
A large scale analysis of cDNA in Arabidopsis thaliana: Generation
of 12,028 non-redundant expressed sequence tags from normalized and
size-selected cDNA libraries
DNA Res. 7 (3), 175-180 (2000)
10907847
Contact: Erika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yama 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.
location/Qualifiers

FEATURES
source 1..525
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/ecotype="Columbia"
/db_xref="taxon:3702"
/clone="APD27d04r"
/tissue_type="aboveground organs"
/dev_stage="two to six-week old"
/clone_lib="Arabidopsis thaliana aboveground organs two to
six-week old"
/note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
XhoI"

ORIGIN

Query Match 44.2%; Score 525; DB 1; Length 525;
Best Local Similarity 100.0%; Pred. NO. 1.3e-142;
Matches 525; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 20 CAACCTTTAATATCTAGTTAATATCAAAATGGCGCGATGAGAGAGTCCAAAGTTT 79

Db 1 CAACCTTTAATATCTAGTTAATATCAAAATGGCGCGATGAGAGAGTCCAAAGTTT 60

Qy 80 CCTCTGTGTAACTCCGGCGGCTTTGAGATGAGAGCTCCGCAACCGAGATACCGG 139

Db 61 CCTGTGTAACTCCGGCGGCTTTGAGATGAGAGCTCCGCAACCGAGATACCGG 120

Qy 140 TGGATCCGCTGGAATAATGATTCAACAGACCCGCAAAACCGGTGAGATCACTCTCCAA 199

Db 121 TGGATCCGCTGGAATAATGATTCAACAGACCCGCAAAACCGGTGAGATCACTCTCCAA 180

Qy 200 CAGTGCAGGAACCTATCCCGTCTTTTATCTTCATGAGCTTTATCTTGGCACTACT 259

Db 181 CAGTGCAGGAACCTATCCCGTCTTTTATCTTCATGAGCTTTATCTTGGCACTACT 240

Qy 260 TCTACTGACGCTTTTAAACAATCGCTTCGATGGTTAACTTTGTAGCCCAAGT 319

Db 241 TCTACTGACGTTCTTAACACATCGCTTCGATGTTAATCTTGTAGACCCACAGT 300

Qy 320 TGTCAAAATATTTGCCCCCGGAGAGGCAATGGAATGAGACATGCTGGAAGTGTATA 379

Db 301 TGTCAAAATATTTGCCCCCGGAGAGGCAATGGAATGAGACATGCTGGAAGTGTATA 360

Qy 380 ACTGGGATCGGAAAACCTCAAAAGCTCACCTACCACTTCGCTTAATGTGTAATAAT 439

Db 361 ACTGGGATCGGAAAACCTCAAAAGCTCACCTACCACTTCGCTTAATGTGTAATAAT 420

Qy 440 ACACCTACTGTGTGGCCCAAGCCGCGTGGAAAACCGCGTTTGGCGTTGCGTTAGGCC 499

Db 421 ACACCTACTGTGTGGCCCAAGCCGCGTGGAAAACCGCGTTTGGCGTTGCGTTAGGCC 480

Qy 500 ATGGCCGACATTTAGACCATCCATCCATGCTTTTACGCTCTTAATG 544

Db 481 ATGGCCGACATTTAGACCATCCATCCATGCTTTTACGCTCTTAATG 525

RESULT 8
DN777869 741 bp mRNA linear EST 05-APR-2005
LOCUS DN777869
DEFINITION E568 [C4091|sltl2.e11] Salinity Library Thellungiella salsauginea
cDNA, mRNA sequence.
ACCESSION DN777869
VERSION DN777869.1 GI:62206750
KEYWORDS EST.
SOURCE Thellungiella salsauginea
ORGANISM Thellungiella salsauginea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Thellungiella.
1 (bases 1 to 741)
Wong, C.E., Li, Y., Whitley, B.R., Diaz-Camino, C., Ahlter, S.R.,
Brandle, J.E., Golding, G.B., Weretilnyk, E.A., Moffatt, B.A. and
Griffith, M.
Expressed sequence tags from the yukon ecotype of Thellungiella
reveal that gene expression in response to cold, drought and
salinity shows little overlap
Plant Mol. Biol. 58 (4), 561-574 (2005)
16021339
Contact: Moffatt, B
Department of Biology
University of Waterloo
200, University Ave West Ontario, Canada N2L 3G1
Tel: (519) 888-4567 ext 2517
Fax: (519) 746-0614
Email: moffatt@uwaterloo.ca.
location/Qualifiers

FEATURES
source 1..741
/organism="Thellungiella salsauginea"
/mol_type="mRNA"
/cultiivar="Yukon"
/db_xref="taxon:72664"
/tissue_type="leave"
/dev_stage="4-week old plants"
/clone_lib="Salinity Library"
/note="Vector: pCMVSPORT6; Site_1: NotI; Site_2: SalI.
Library was prepared following the protocol for Invitrogen
SuperScript Plasmid System with Gateway Technology for
cDNA synthesis and cloning"

ORIGIN

Query Match 44.1%; Score 524; DB 8; Length 741;
Best Local Similarity 85.5%; Pred. NO. 2.7e-142;
Matches 595; Conservative 0; Mismatches 100; Indels 1; Gaps 1;

Qy 39 TTTAATCAATATGCGCGGATGAGAGACATGCAAGCTTTCTCTGTGTAACTCCGCGC 98

Db 14 TATCAAAAATATGCGCGGATGAGAGACATGCAAGCTTTCTCTGTGTAACTCCGCGC 73

Qy 99 GCGTTTGAATAGACGCTCCCGCAACCGAGATACCGGTGATCCGATGGAATAATGA 158

Db 74 GGGCTTTGAGAGGACGCTTCCGACCAACCGGTATCAACGGCGATCTTGGAAACCGA 133

Qy 159 TTCAACAGACCGCCAAACCGGTGAGATCACTGTCCAAACGTGGCGAACTTATCC 218

Db 134 TTCCGACGCGCGCTTAAGCGGTGAGATCACTCCCAACAGTTGCCGAACTTACC 193

Qy 219 CGTGGTTTATCTTCCATGCTTTTATCTTCCGAAACATCTTCACTCTGAAGTTCTTAA 278

Db 194 CGTGGTTTATCTTCCATGATTTTATCTTCCGAAACATCTTCACTCTGAAGTTCTTAA 253

Qy 279 CCACATCGCTTCGATGCTTACATCTTGTAGCCCAAGTTGCAATATTATTCGCCCC 338

Db 254 CCAAGTGGCTTCTCATGCTCATCTTGTAGCCCAAGTTGCAATATTATTCGCCCC 313

Qy 339 GGGAGGGCAATGAGATGAGACGATGCTGGAAGTGTGATTAACCTGGGCGATCGAAAACTT 398

Db 314 GGGAGGGCAATGAGATGAGACGATGCTGGAAGTGTGATTAACCTGGGCGATCGAAAACTT 373

Qy 399 CAAAGCTACCTACCACTTGGTAAATGCTAAATGAAATAACCTCACTCGTGGGCGA 458

Db 374 CAAAGCTACCTCCCAACTCGTAAACGCTAAATGAAATAACCTCACTCGTGGGCGA 433

Qy 459 CAGCGCGGTGGGAAACCGCGTTGCGGTTGCGCTAGGCGCATCCGCAATTAGACC 518

Db 434 TAGCGGGGTGGCAAAACCGGTTGCGGTTGCTTTAGGCCATCCGCAATTAGACC 493

Qy 519 ATCATCACTTTTCACTTAAATGAAATGATTCAGTGGCAGAACTTAACTAAATCAT 578

Db 494 ATCTTCAATTTTCACTTCAATGAAATGACCCGGTAGCAAGAACTCACTAAATCAT 553

Qy 579 TAGAAGCGATCGGATATCTTAAGTAAACCGGAATCTTTCAGCTGGGCAATACCGGT 638

Db 554 GAGAACCGATCGGAAATCTTAACGTAATTAACCGGAATCTTTCAGTGGGCAATACCGGT 613

Qy 639 TGCAGTGGTGGAAACCGGACTCGAACCGAATGAGAACATCTGATGCAATCGCAACC 698

Db 614 TGCAGTGGTGGGACCGGCTCTCGAACCGAATGAGAACATCTGATGCAATCGCAACC 673

Qy 699 AACGG-ACCTTAACCATGAGAGTTTAACTAAAGAT 733

Db 674 AGCGAAGTGAACCATGAGAGTTTAAAGATGT 709

RESULT 9
B20290/c
LOCUS 794 bp DNA linear GSS 16-SEP-1997
DEFINITION T21K8-T7 TAMU Arabidopsis thaliana genomic clone T21K8, genomic
survey sequence.
ACCESSION B20290
VERSION B20290.1 GI:2395344
KEYWORDS GSS.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
REFERENCE Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
1 (bases 1 to 794)
Feng, J., Dewar, K., Buehler, E., Kim, C., Li, Y., Shinn, P., Sun, H. and
Ecker, J.
TITLE BAC End Sequences at ATGC
JOURNAL Unpublished (1997)
COMMENT Other_GSSs: T21K8-Sp6
Contact: Ecker J.
Arabidopsis Thaliana Genome Center
University of Pennsylvania
Dept. of Biology, University of Pennsylvania, Philadelphia, PA
19104
Tel: 215-898-9384
Fax: 215-898-8780
Email: jaecker@genome.bio.upenn.edu
Seq primer: T7
Class: BAC ends
High quality sequence start: 100

FEATURES
source
High quality sequence stop: 448.
Location/Qualifiers
1..794
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/ecotype="Columbia"
/db_xref="taxon:3702"
/clone="T21K8"
/sex="hermaphrodite"
/note="Vector: Belobacii; Site 1: HindIII; Site 2:
HindIII; Produced by Rod Wing"

Query Match 42.6%; Score 506.6; DB 9; Length 794;
Best Local Similarity 90.0%; Pred. No. 3.6e-137;
Matches 561; Conservative 0; Mismatches 60; Indels 2; Gaps 2;

Qy 358 GAGATGCTGGAAGTGTGATTAACCTGGGCAATCGGAAACCTCAAGCTCACTACCACT 417

Db 693 GAGAAATGTGAAAGTGTGATTAAGTGAATTTGGGAAACCTGGAACCTTACATACCACT 634

Qy 418 TCGGTAATGCTAATGGAATAATACCTCACTGCGGCGACAGCGCGGTGGGAAACG 477

Db 633 TCGGTAATGCTAATGGAATAATACCTCACTGCGGCGACAGCGCGGTGGGAAACG 574

Qy 478 GCGTTTGGGTTGCGCTAGGCGCATGCGCAACATTAAGCCATCACTACGTTTTCAGCT 537

Db 573 GCGTTTGGGTTGCGCTAGGCGCATGCGCAACATTAAGCCATCACTACGTTTTCAGCT 514

Qy 538 CTAAATGGAATTTGATTCAGTGGGCAAGCACTTAACCAATCACTTGAACCAATCCGCAATTC 597

Db 513 TTAAATGGAATTTGATTCAGTGGGCAAGCACTTAACCAATCACTTGAACCAATTC 454

Qy 598 TTAACGTAATTAACCGGAATCTTTCAGTGGGCAATCACTACCGGTGAGTGGGAAACCGGA 657

Db 453 TTAACGTAATTAACCGGAATCTTTCAGTGGGCAATCACTACCGGTGAGTGGGAAACCGGA 395

Qy 658 CTCGACCGAAGTGAACCAAGTATGCCACATTCGCGACCAACGGAATTTAAACATGAG 717

Db 394 CTCGACCGAAGTGAACCAAGTATGCCACATTCGCGACCAACGGAATTTAAACATGAG 335

Qy 718 GAGTTTTCAAAGAGTGAAGGCGAGCAAGGCCATTTTGGCTCGGATTCGCAAT 777

Db 334 GAGTTTTCAAAGAGTGAAGGCGAGCAAGGCCATTTTGGCTCGGATTCGCAAT 275

Qy 778 ATGATATATGTTGACGATGATTTGCGCGGTTTGGTTTAAAGCCGCTGATGATGT 837

Db 274 ATGATATATGTTGACGATGATTTGCGCGGTTTGGTTTAAAGCCGCTGATGATGT 215

Qy 838 AAGATGCGCAAGAAAGAAAGTGTGATGAGAGCTTTGATGATGATGATGATGATG 897

Db 214 AAGATGCGCAAGAAAGAAAGTGTGATGAGAGCTTTGATGATGATGATGATGATG 155

Qy 898 TTTCTCAAGTATGTTTGGGCGTGAAGAAAGCGGAGATTCATGATGATGATGATGATG 957

Db 154 TTTCTCAAGTATGTTTGGGCGTGAAGAAAGCGGAGATTCATGATGATGATGATGATG 95

Qy 958 TCCG-TTTCCTCGGCGCAAGCTTG 979

Db 94 TCCGTTTATCCGCGCAAGCTTG 72

RESULT 10
A1997409/c
LOCUS 547 bp mRNA linear EST 08-SEP-1999
DEFINITION 701553812 A. thaliana, Columbia Col-0, root-2 Arabidopsis thaliana
CDNA clone 701553812, mRNA sequence.
ACCESSION A1997409
VERSION A1997409.1 GI:5844314
KEYWORDS EST.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana

REFERENCE
AUTHORS

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi-
1 (bases 1 to 547)
Chen, J., Momiyama, M., Chan, E., Mooney, M., Carroon, B., Gilliland, D., Wang, X., Hillman, J., Guegler, K., Kim, C., Doyle, M., Brzoska, P., Gorgone, G., Burns, D., Griffin, J., Mouanoutou, M., Nguyen, D., Tan, R., Rose, M., Warren, B., Ton, B., Kastury, K., Borillo, C., Cardio, T., Policky, J., Suzuki, G., Argentine, C., Shah, S., Nobrigo, A., Murry, L., Turner, C., Krikorian, S., Elder, L. and Hanson, D.

TITLE
JOURNAL
COMMENT

Arabidopsis thaliana Gene Expression MicroArray
Unpublished (1999)
Contact: David Smoller, Ph.D.
Genome Systems, Inc., a wholly owned subsidiary of Incyte Pharmaceuticals, Inc.,
463 World Parkway Circle, St. Louis, MO 63134, USA
Tel: 877-577-2733
Fax: 314-427-3324
Email: service@genomesystems.com.

FEATURES
SOURCE

1..547
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/ecotype="Col-0"
/db_xref="taxon:3702"
/clone="701553812"
/tissue_type="root"
/dev_stage="4 - 7 weeks"
/clone_lib="A. thaliana, Columbia Col-0, root-2"
/note="Vector: pSPORT; Site 1: NotI; Site 2: SalI; cDNA library was derived from untreated root tissue from Arabidopsis thaliana, Columbia Col-0, at 4 - 7 weeks. Plants were grown in 1:1:1 peat moss/vermiculite/perlite soil at 22 deg. C +/- 3 deg. C under constant light, and watered with fertilizer. cDNA synthesis was initiated using a NotI-oligo(dT) primer. Double-stranded cDNA was bluntended, ligated to SalI adaptors, digested with NotI, size-selected, and cloned into the NotI and SalI sites of the pSPORT vector."

ORIGIN

Query Match 41.6%; Score 493.8; DB 1; Length 547;
Best Local Similarity 94.2%; Pred. No. 1.9e-133;
Matches 517; Conservative 0; Mismatches 30; Indels 2; Gaps 2;

QY 550 CGCATATCTTAACGATTAACCGGAATCTTTCGAGCTGACATACCGGTTGCAGTGGG 649
Db 547 CGCATATCTTAACGATTAACCGGA-CTTTCGAGCT-GACATACCGGTTGCAGTGGG 490
QY 650 GAACCGGACTCGGACCGAAGTGAACAAGTATGCCATGCGGACCAACGGAAGCTTAA 709
Db 489 NAACCGGACTCGGACCGAAGTGAACAAGTATGCCATGCGGACCAACGGAAGCTTAA 430
QY 710 ACCATGAGAGTTTAAACAAGTGTAAAGGACGAAAGCCCATTTCTGTGCTCGGAGTT 769
Db 429 ACCATGAGAGTTTAAACAAGTGTAAAGGACGAAAGCCCATTTCTGTGCTCGGAGTT 370
QY 770 ACGGACATATGATATGTTGACGATGATTTGCCCGTTTGTGGGTTTATGGCCGTT 829
Db 369 ACGGACATATGATATGTTGACGATGATTTGCCCGTTTGTGGGTTTATGGCCGTT 310
QY 830 GTATGTATGATGAGTGGCAAGAAAAGTCTAGATGAGGAGCTTTTAGTGAATG 889
Db 309 GTATGTATGATGAGTGGCAAGAAAAGTCTAGATGAGGAGCTTTTAGTGAATG 250
QY 890 TGGTTCGTTTCTCAAGTATAGTTTGTGGGTTGAAAAAGCGAAGTTGATTTGATGGA 949
Db 249 TGGTTCGTTTCTCAAGTATAGTTTGTGGGTTGAAAAAGCGAAGTTGATTTGATGGA 190
QY 950 AGGATCTCTCCGTTTCTCCGCGCAAGCTTGAATCTTCACTGAGTTGGAAGAAGCTTCTG 1009
Db 189 AGGATCTCTCCGTTTCTCCGCGCAAGCTTGAATCTTCACTGAGTTGGAAGAAGCTTCTG 130

QY 1010 GTATCTTCGTCATGATTGTGTTATGATATATATGACAGAGGGCTTGAATTTGAAA 1069
Db 129 GTATCTTCGTCATGATTGTGTTATGATATATGACAGAGGGCTTGAATTTGAAA 70

QY 1070 ACCATCAATGTTTTCAGCTCCAGCTAGCTATTTTCATGTCCTTAAGTTGATGTA 1129
Db 69 ACCATCAATGTTTTCAGCTCCAGCTAGCTATTTTCATGTCCTTAAGTTGATGTA 10

QY 1130 TTTTATTATA 1138
Db 9 TTTTATTATA 1

RESULT 11

AV561535 413 bp mRNA linear EST 23-FEB-2004
LOCUS AV561535 Arabidopsis thaliana green siliques Columbia Arabidopsis
DEFINITION thaliana cDNA clone SQ153b01P 3', mRNA sequence.
ACCESSION AV561535
VERSION AV561535.1 GI:8732961
KEYWORDS EST.

SOURCE
ORGANISM

Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi-
1 (bases 1 to 413)

REFERENCE
AUTHORS
TITLE

Asamizu, B., Nakamura, Y., Sato, S. and Tabata, S.
A large scale analysis of cDNA in Arabidopsis thaliana: Generation of 12,028 non-redundant expressed sequence tags from normalized and size-selected cDNA libraries
DNA Res 7 (3), 175-180 (2000)

JOURNAL
PUBMED

10907847
Contact: Erika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizukazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.

FEATURES
SOURCE

1..413
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/ecotype="Columbia"
/db_xref="taxon:3702"
/clone="SQ153b01P"
/tissue_type="green siliques"
/clone_lib="Arabidopsis thaliana green siliques Columbia"
/note="Vector: pBluescriptII SK-, site_1: EcoRI; site_2: XhoI"

ORIGIN

Query Match 34.8%; Score 413; DB 1; Length 413;
Best Local Similarity 100.0%; Pred. No. 1e-109;
Matches 413; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 533 CAGCTCAATATGAAATGATCCAGTCGACAGAACTAACAATCATATAGAACGATTCGC 592
Db 1 CAGCTCAATATGAAATGATCCAGTCGACAGAACTAACAATCATATAGAACGATTCGC 60
QY 593 ATATCTTAACGATTAACCGGAATCTTTCAGCTGACATACCGGTTGAGTGGGAA 652
Db 61 ATATCTTAACGATTAACCGGAATCTTTCAGCTGACATACCGGTTGAGTGGGAA 120
QY 653 CCGGACTCGGACCGAAGTGAACAAGTGTATGCCATATCGGACCAACGGAAGCTTAAAC 712
Db 121 CCGGACTCGGACCGAAGTGAACAAGTGTATGCCATATCGGACCAACGGAAGCTTAAAC 180
QY 713 ATGAGAGTTTAAACAAGTGTATGAGGCGAAGGCCATTCGTTGCGGATTAACG 772
Db 181 ATGAGAGTTTAAACAAGTGTATGAGGCGAAGGCCATTCGTTGCGGATTAACG 240
QY 773 GACATATGATATGTTGACGATGATTTGCCCGTTTGTGGGTTATGCGCGTTGTA 832

Db 241 GACATATGATATGTTGACCATGATTTGCCGGTTTGTGGTTTATGGCCGGTTTA 300

Qy 833 TGTGTAAAGATGGCGCAAGAAAAAGTGTAGATAGAGAGCTTTGTAGTGGAAATTGGG 892

Db 301 TGTGTAAAGATGGCGCAAGAAAAAGTGTAGATAGAGAGCTTTGTAGTGGAAATTGGG 360

Qy 893 TTGCGTTCTCAAGATATGTTGTGGGGTGAAGAAAGCGGAGATTGATTGATT 945

Db 361 TTGCGTTCTCAAGATATGTTGTGGGGTGAAGAAAGCGGAGATTGATTGATT 413

RESULT 12
DN779181/c 570 bp mRNA linear EST 05-APR-2005

LOCUS DN779181 E756 [C4680]srl_f12 salinity subtracted library Thellungiella

DEFINITION DN779181 salunginea cDNA, mRNA sequence.

ACCESSION DN779181

VERSION DN779181.1 GI:62208062

KEYWORDS EST

SOURCE Thellungiella salunginea

ORGANISM Thellungiella salunginea; Streptophyta; Embryophyta; Tracheophyta; Eukaryota; Viridiplantae; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Thellungiella.

REFERENCE 1 (bases 1 to 570) Wong,C.E., Li,Y., Whitty,B.R., Diaz-Camino,C., Akhter,S.R., Brandie,J.E., Golding,G.B., Weretilnyk,E.A., Moffatt,B.A. and Griffith,M.

AUTHORS Expressed sequence tags from the Yukon ecotype of Thellungiella reveal that gene expression in response to cold, drought and salinity shows little overlap

JOURNAL Plant Mol. Biol. 58 (4), 561-574 (2005)

COMMENT 16021339

CONTACT: Moffatt,B

DEPARTMENT OF BIOLOGY

UNIVERSITY OF WATERLOO

200, University Ave West, Ontario, Canada N2L 3G1

TEL: (519) 888-4567 ext 2517

FAX: (519) 746-0614

EMAIL: moffatt@uwaterloo.ca.

LOCATION/Qualifiers

FEATURES

source 1..570

/organism="Thellungiella salunginea"

/mol_type="mRNA"

/cultivar="Yukon"

/db_xref="taxon:72664"

/tissue_type="leaf"

/dev_stage="4-week old plants"

/clone_lib="Salinity Subtracted Library"

/note="Vector: pCMVSPORT6; Site 1: NotI; Site 2: NotI; Library was prepared following the protocol for BD Clontech PCR-Select cDNA Subtraction Kit and Promega pGEM-T Easy Vector System for cDNA synthesis and cloning"

ORIGIN

Query Match 34.7%; Score 412.4; DB 8; Length 570;

Best Local Similarity 85.0%; Pred. No. 1.6e-109;

Matches 472; Conservative 0; Mismatches 82; Indels 1; Gaps 1;

Qy 478 GCGTTGGGGTGGGCTAGGCGGCAATTAAGCCATCAGCTTTTCAGCT 537

Db 570 GCGTTGGGGTGGGCTTAGGCGGCAATTAAGCATATCC-TCAANTTTTCAGCT 512

Qy 538 CTAATAGGAATGATTCAGATCGAGAACTAACAAATCATTAGAACGATCCGATATC 597

Db 511 CTCAATAGGAATCGAACCGGTAGCAGGAATGCAAAATGACAGAACGATCCGGAATC 452

Qy 598 TTAACGTTAAACCGGAATCTTTGAGCTGCAATACCGGTTGCAAGTGGTGGAAACCGGA 657

Db 451 TTAACGTTAAACCGGAATCTTTGAGCTGCAATACCGGTTGCAAGTGGTGGAAACCGGT 392

Qy 658 CTGGAACGGAAGTGAACAGGTATGACCATGCGGACCAAGGACTTAAACATGAG 717

Db 391 CTGGACCCAGAGATTAACGGGATGATATCCACCATGCGCACCGAGGAAGTGAACATGAG 332

Qy 718 GAGTTTAAACAAGATGTAAGGCGACGAAAGCCATTTCGTGGCTGCGATTAACGACAT 777

Db 331 GAGTTTAAACAAGATGTAAGGCGACGAAAGCCATTTCGTGGCTGCGATTAACGACAT 272

Qy 778 ATGATATGTTGACGATGATTTGCCGGTTTGTGGTTTATGGCCGGTTGATGTGT 837

Db 271 ATGATATGTTGACGATGATTTGCCGGTTTGTGGTTTATGGCCGGTTGATGTGT 212

Qy 838 AAGAATGGCAAGAAAAAGTCTGAGATAGAGAGCTTTAGTGGTAATTGCGTTCG 897

Db 211 AAGAATGGCAAGAAAAAGTCTGAGATAGAGAGCTTTAGTGGTAATTGCGTTCG 152

Qy 898 TTTCTCAAGATATGTTGTGGGGTGAAGAAAGCGGAGATTGATTTGGAAGATCCT 957

Db 151 TTTCTCAAGATATGTTGTGGGGTGAAGAAAGCGGAGATTGATTTGGAAGATCCT 92

Qy 958 TCGTTTCTCGGCGCAAGCTTATCTTCACTGAGTTGGAAGAGCTTCTGTATCTTC 1017

Db 91 TCGCTCTCGGCGGAGCGTGTATCTTCCCGAGCTGGAAGAGCTTCTCGGCTACTTC 32

Qy 1018 GTCATGATTTGTGTT 1032

Db 31 GTCATGATTTGTGTT 17

RESULT 13
DN779131/c 547 bp mRNA linear EST 05-APR-2005

LOCUS DN779131 E7515 [C4629]srl_b09 salinity subtracted library Thellungiella

DEFINITION DN779131 salunginea cDNA, mRNA sequence.

ACCESSION DN779131

VERSION DN779131.1 GI:62208012

KEYWORDS EST

SOURCE Thellungiella salunginea

ORGANISM Thellungiella salunginea; Streptophyta; Embryophyta; Tracheophyta; Eukaryota; Viridiplantae; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Thellungiella.

REFERENCE 1 (bases 1 to 547) Wong,C.E., Li,Y., Whitty,B.R., Diaz-Camino,C., Akhter,S.R., Brandie,J.E., Golding,G.B., Weretilnyk,E.A., Moffatt,B.A. and Griffith,M.

AUTHORS Expressed sequence tags from the Yukon ecotype of Thellungiella reveal that gene expression in response to cold, drought and salinity shows little overlap

JOURNAL Plant Mol. Biol. 58 (4), 561-574 (2005)

COMMENT 16021339

CONTACT: Moffatt,B

DEPARTMENT OF BIOLOGY

UNIVERSITY OF WATERLOO

200, University Ave West, Ontario, Canada N2L 3G1

TEL: (519) 888-4567 ext 2517

FAX: (519) 746-0614

EMAIL: moffatt@uwaterloo.ca.

LOCATION/Qualifiers

FEATURES

source 1..547

/organism="Thellungiella salunginea"

/mol_type="mRNA"

/cultivar="Yukon"

/db_xref="taxon:72664"

/tissue_type="leaf"

/dev_stage="4-week old plants"

/clone_lib="Salinity Subtracted Library"

/note="Vector: pCMVSPORT6; Site 1: NotI; Site 2: NotI; Library was prepared following the protocol for BD Clontech PCR-Select cDNA Subtraction Kit and Promega pGEM-T Easy Vector System for cDNA synthesis and cloning"

ORIGIN

Query Match 33.9%; Score 403; DB 8; Length 547;

Best Local Similarity 84.9%; Pred. No. 9.5e-107;
Matches 451; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

Qy	502	ACCGCAACATTAGAACCCATCCATCCATCAGCTTTTACGCTCTAATATGGAATTGATCACTCGCA	561
Db	547	GCCGCACATTAGACACACATCTTTCMAATTTTTCAGCTCTCATAGGAATGACCCGGTAGCA	488
Qy	562	GGAACTTAACAATAATCATTTAGAACCGGATCCGCAATATCTTAATGATTAACCGGAATCTTTC	621
Db	487	GGAAATCAGCAAAATGATGATAGAACCCATCCGGAAATCTTAACGATTAACCGGAATCTTTC	428
Qy	622	GAGCTGCACTACCCGCTTGCACTGGTGGGAACCGGACTCGACCGGAAGTGAACAACGCG	681
Db	427	GAGATAGACATACCGGTTGCACTGATGATCGGAGCGGCTTCGAGACCGAAGTACCGGATG	368
Qy	682	ATGCCACCATGCGCACCAACGGACTTAAACCATAGAGATTTTACAAAGATGTAAAGCG	741
Db	367	ATGCCACCATGCGCACACCGGAATGAAACCATAGAGATTTTAAAGATGTAAAGCGC	308
Qy	742	ACGAAAGCCCATTCCTGCTGGCTGCGGATTAACGGAATATGGAATTTGAGACGATTTG	801
Db	307	ACCAAGGACATTTGCTGAGCTGCTGATTAATGACATATGGAATATTTGACATATTTG	248
Qy	802	CCCGGTTTGTGGGTTTATGGCCGGTGTATGTATGTAAGATGGCAAGAAAAAGCT	861
Db	247	CCCGGTTTGTGGGTTTATGGCCGGTGTATGTATGTAAGAACGGAAAAAGTAAAGT	188
Qy	862	GAGATGAGAGACTTTGTAGTGGGAATTTGTGTTGGGCTTCTCAATATATGTTTGGGGT	921
Db	187	GAGATGAGAGAAATTTGTGAGGATATTTGTGATTTCTCAAGATATATGTTTAAAGGCT	128
Qy	922	GAAGAAAGCGAGATTCGATTTGATTTGGAAGATCCCTCCGTTCTCCGCGCAAGCTTGAT	981
Db	127	CACAAATCGAGATTCGAGAGATTCGAAAGGATCTCTTGCGCTCCCGGTGACGGTTGAT	68
Qy	982	CCTTCACTGAGTTGGAAGAAAGCTTCTGTATCTTGCTCTAGATTTGTGCTT	1032
Db	67	CCTTTGCCCGAGCTGAAGAGGCTTCGGGCTACCTTCGCTTAATTTGGTCT	17

```

/moi_type="mRNA"
/db_xref="taxon:3702"
/clones="RAFL16-6c-J07"
/lab_host="DH10B"
/clone_id="RAFL16"
/note="Site_1: BamHI; Site_2: SalI; dark-green"

```

Query Match	Similarity	Score	DB	Length
Best Local	99.0%	Pred. No. 2e-103;		430;
Matches	403;	Conservative	0;	Mismatches 3; Indels 1; Gaps 1;
Qy	766	GATTACGACATATGATATGTTGGACGATATTCGCCGGTTTGTGTGGTTATGACC	825	
Db	407	GATTACCGACATAT-GATATGTTGGACGATATTCGCCGGTTTGTGTGGTTATGACC	349	
Qy	826	GGTTGTATGCTATGAATGGGCCAAAGAAAAAGCTGAGATGAGGACCTTGTATGTTGA	885	
Db	348	GGTTGTATGCTATGAATGGGCCAAAGAAAAAGCTGAGATGAGGACCTTGTATGTTGA	289	
Qy	886	ATTGTGGTGGCTTCTCAAGTATAGTTTGGGGGTGA AAAAGCGGAGATTCGATTATT	945	
Db	288	ATTGTGGTGGCTTCTCAAGTATAGTTTGGGGGTGA AAAAGCGGAGATTCGATTATT	229	
Qy	946	GTCGAAGATCTTCGCTTCTCCGGCCAAAGCTTATCTTCACTGATGGAGAAAGCT	1005	
Db	228	GTCGAAGATCTTCGCTTCTCCGGCCAAAGCTTATCTTCACTGATGGAGAAAGCT	169	
Qy	1006	TCGTGATCTTCGCTATGATTTGGTATGACTATATTCAGAGGGGCTTGATATTGG	1065	
Db	168	TCGTGATCTTCGCTATGATTTGGTATGACTATATTCAGAGGGGCTTGATATTGG	109	
Qy	1066	AAAAACCTATCAATGTTTTCTAGCTCCAGCTAGCTATATGTCATGTCCTAAGTTCATG	1125	
Db	108	AAAAACCTATCAATGTTTTCTAGCTCCAGCTAGCTATATGTCATGTCCTAAGTTCATG	49	
Qy	1126	TGTAATTTTATTAATCGATCAACAATTTGTATATGTTTATACC	1172	
Db	48	TGTAATTTTATTAATCGATCAACAATTTGTATATGTTTATACC	2	

site (http://www.gsc.riken.go.jp/e/plant/index_e.html) for further details.

FEATURES

Source

Location/Qualifiers

1. 411

/organism="Arabidopsis thaliana"

/mol_type="mRNA"

/db_xref="taxon:3702"

/clone="RAFL09-95-F15"

/dev_stage="plants at various developmental stages from germination to mature seeds"

/lab_host="DH10B"

/clone_id="RAFL9"

/note="Site 1: BamHI; Site 2: SalI; subjected to dehydration (1, 2, 5, 10, 24 hr) and cold (1, 2, 5, 10, 24 hr) treatments"

ORIGIN

Query Match 32.8%; Score 389.4; DB 1; Length 411;

Best Local Similarity 99.7%; Pred. No. 8.8e-103;

Matches 390; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 786 GTTGACGATGATTTGCCGGTTTGTGGGTTTAATGCGCGTTGTATGTGAAGATG 845
Db 411 GTTGACGATGATTTGCCGGTTTGTGGGTTTAATGCGCGTTGTATGTGAAGATG 352
QY 846 GCAAGAGAAAAAGCTGATGAGAGCTTTGTAGTGAATTTGTGGCTTTCTCA 905
Db 351 GCAAGAGAAAAAGCTGATGAGAGCTTTGTAGTGAATTTGTGGCTTTCTCA 292
QY 906 GTATAGTTTGTGGGTGAAAAAGCGAGATTCGATTGATGTGAAGATCCTTCGTTTC 965
Db 291 GTATAGTTTGTGGGTGAAAAAGCGAGATTCGATTGATGTGAAGATCCTTCGTTTC 232
QY 966 TCCGGCCAAAGCTTGAATCCTTCACTGAGTGAAGAACTTCTGTATCTTCTAGAT 1025
Db 231 TCCGGCCAAAGCTTGAATCCTTCACTGAGTGAAGAACTTCTGTATCTTCTAGAT 172
QY 1026 TTGTGTATGATATATTCAGAGGGGCTTGAATATTGAAAAAACCATCATGTTTC 1085
Db 171 TTGTGTATGATATATTCAGAGGGGCTTGAATATTGAAAAAACCATCATGTTTC 112
QY 1086 TAGCTCCAAAGCTAGTATTTCAATGCTCAAGTTGATGTAATTTTATTAACTGA 1145
Db 111 TAGCTCCAAAGCTAGTATTTCAATGCTCAAGTTGATGTAATTTTATTAACTGA 52
QY 1146 TCAAAACATTTGTTATAGTTTACCCAAA 1176
Db 51 TCAAAACATTTGTTATAGTTTACCCAAA 21
```

Search completed: March 20, 2006, 16:52:58
Job time : 5388.1 secs

THIS PAGE LEFT BLANK

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 20, 2006, 13:51:15 ; Search time 699.094 Seconds

(Without alignments)
11325.590 Million cell updates/sec

Title: US-10-634-548-18

Perfect score: 1188
Sequence: 1 gatcataaattctcaacac.....ccccaaaaaaaaaaaaa 1188

Scoring table: IDENTITY_NUC
Gapex 10.0 , Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_21:*

- 1: geneseqn1980s:*
- 2: geneseqn1990s:*
- 3: geneseqn2000s:*
- 4: geneseqn2001as:*
- 5: geneseqn2001bs:*
- 6: geneseqn2002as:*
- 7: geneseqn2002bs:*
- 8: geneseqn2003as:*
- 9: geneseqn2003bs:*
- 10: geneseqn2003cs:*
- 11: geneseqn2003ds:*
- 12: geneseqn2004as:*
- 13: geneseqn2004bs:*
- 14: geneseqn2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1188	100.0	1188	5 AAH41136	Aah41136 Arabidops
2	1188	100.0	1188	12 ADJ98180	Adj98180 Thale cre
3	1188	100.0	1188	3 AAC34054	Aac34054 Arabidops
4	975	82.1	975	6 ABZ12353	ABZ12353 Arabidops
5	975	82.1	975	6 ADG87791	ADG87791 A. thalia
6	975	82.1	975	6 ADG87590	ADG87590 A. thalia
7	748	63.0	1125	13 ADX51602	ADX51602 Plant ful
8	239.4	20.2	987	6 AAD38074	Aad38074 Grape chl
9	202.6	17.1	1174	6 AAD38077	Aad38077 Soybean c
10	199	16.8	1242	6 AAD38082	Aad38082 Wheat chl
11	194.4	16.4	1104	6 AAD38078	Aad38078 Soybean c
12	188	15.8	1125	6 AAD38079	Aad38079 Soybean c
13	178.8	15.1	1444	6 AAD38080	Aad38080 Tulip chl
14	169.8	14.3	1215	6 AAD38075	Aad38075 Grape chl
15	163.4	13.8	1135	5 AAH41137	Aah41137 Arabidops
16	163.4	13.8	1135	12 ADJ98181	Adj98181 Thale cre
17	151.6	9.3	1340	5 AAH41135	Aah41135 Chenopodi
18	110.8	9.3	1156	6 AAD38086	Aad38086 Wheat chl
19	109.2	9.2	1302	6 AAD38076	Aad38076 Corn chlo

ALIGNMENTS

20	106.6	9.0	1244	13 ADX50647	Adx50647 plant ful
21	106.6	9.0	1244	13 ADX50646	Adx50646 plant ful
22	98.6	8.3	916	6 AAD38081	Aad38081 wheat chl
23	82.6	7.0	1274	6 AAD38087	Aad38087 corn chlo
24	67	5.6	1300	5 AAH41145	Aah41145 chenopodi
25	63.8	5.4	558	11 ACU32937	Acu32937 rice abio
26	53	4.5	390	11 ACU27601	Acu27601 rice abio
27	47	4.0	985	6 ABZ15053	Abz15053 arabidops
28	45	3.8	1829	2 AAX84982	Aax84982 human sec
29	45	3.8	1829	8 ACD18908	Acid18908 novel hum
30	45	3.8	1829	12 ADG78299	Adg78299 human sec
31	45	3.8	1829	12 ADN60590	Adn60590 human sec
32	44.8	3.8	2000	8 ADA17938	Ada17938 rice gene
33	41.6	3.5	631	5 ADI73251	Adi73251 human ova
34	41.6	3.5	631	5 ADL38382	Adl38382 human ova
35	41.6	3.5	2000	8 ADA17938	Ada17938 rice gene
36	40.8	3.4	13376	6 ABL32583	Ab132583 human imm
37	40.4	3.4	6715	13 ADR84259	Adr84259 aspergill
38	39.8	3.4	752	5 ADL44766	Adl44766 human ova
39	39.6	3.3	604	4 AAH34528	Aah34528 human col
40	39.6	3.3	1345	12 ADQ22878	Adq22878 human sof
41	39.6	3.3	2000	11 ACU35601	Acu35601 rice stre
42	39.6	3.3	4419	6 ABS53449	Abes53449 cdna enco
43	39.4	3.3	452	13 ACF83451	Acf83451 human sir
44	39.3	3.3	2000	11 ACU35887	Acu35887 rice stre
45	38.8	3.3	2000	11 ACU35887	Acu35887 rice stre

RESULT 1

AAH41136 ID AAH41136 standard; DNA; 1188 BP.

AC AAH41136;

DT 21-AUG-2001 (first entry)

DE Arabidopsis thaliana chlorophyllase coding sequence #1.

KM Chlorophyllase; transgenic plant; ds.

OS Arabidopsis thaliana.

XX JP2001086990-A.

PD 03-APR-2001.

PF 20-SEP-1999; 99JP-00266181.

PR 20-SEP-1999; 99JP-00266181.

PA (KAGO) KAGOME KK.

DR WPI, 2001-338421/36.

DR P-PSDB; AAB99102.

PT DNA encoding chlorophyllase, useful for producing transgenic plants.

PS Claim 3; Page 10-11; 21pp; Japanese.

CC The present sequence is a chlorophyllase protein coding sequence. The chlorophyllase protein coding sequence can be used for the transformation of a plant

XX Sequence 1188 BP; 327 A; 255 C; 281 G; 325 T; 0 U; 0 Other;

Query Match 100.0%; Score 1188; DB 5; Length 1188;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GATCATTAATCTTCAACAACTCTTAATTATCTTATTATACAAATGGCGGCGATA 60

```

Db      1 GATCATTAATCTTCAACAACTCTTAATATCTACTTTAATACAAATGGCGCGATA 60
Qy      61 GAGGACAGTCCAGCGTTTCTCTGTGTAACTCCGCGCGCTTTGAGATAGCGACCTC 120
Db      61 GAGGACAGTCCAGCGTTTCTCTGTGTAACTCCGCGCGCTTTGAGATAGCGACCTC 120
Qy      121 CCGACAAACGAGATACCGGTGGATCCGGTGGAAAAATGATCAACGACCGCAAAAACG 180
Db      121 CCGACAAACGAGATACCGGTGGATCCGGTGGAAAAATGATCAACGACCGCAAAAACG 180
Qy      181 GTGAGAAATCACCTGTCCAAACAGTCGCGGAACTTAATCCCGTGTATTTCTTCATGCG 240
Db      181 GTGAGAAATCACCTGTCCAAACAGTCGCGGAACTTAATCCCGTGTATTTCTTCATGCG 240
Qy      241 TTTTATCTTGGCACTACTCTGATCTGACGTTCTTAACCAATCGCTTCCGATGCTTAC 300
Db      241 TTTTATCTTGGCACTACTCTGATCTGACGTTCTTAACCAATCGCTTCCGATGCTTAC 300
Qy      301 ATTCTGTAGCCCAAGTTGTGCAAAATTTAGCCGCGGAGAGGCAAGTGAAGTGGAC 360
Db      301 ATTCTGTAGCCCAAGTTGTGCAAAATTTAGCCGCGGAGAGGCAAGTGAAGTGGAC 360
Qy      361 GATGCTGAGAGTGTGATTAATCTGAGCATCGGAAAACTCAAAAGCTCACTCACTTAC 420
Db      361 GATGCTGAGAGTGTGATTAATCTGAGCATCGGAAAACTCAAAAGCTCACTCACTTAC 420
Qy      421 GTAAATGCTAATGAAAAATACACCTCACTCGTGGGCGCACGCGGGTGGAAAAACGGCG 480
Db      421 GTAAATGCTAATGAAAAATACACCTCACTCGTGGGCGCACGCGGGTGGAAAAACGGCG 480
Qy      481 TTGGCGTGTGGCGCTAGGCGCATGCGGCAACATTAGACCATCATCACTTTCAGCTCTA 540
Db      481 TTGGCGTGTGGCGCTAGGCGCATGCGGCAACATTAGACCATCATCACTTTCAGCTCTA 540
Qy      541 ATAGGAATTTGATCCAGTCGAGGAACTTAACAAATTCATTAGAACCGATCCGATATCTTA 600
Db      541 ATAGGAATTTGATCCAGTCGAGGAACTTAACAAATTCATTAGAACCGATCCGATATCTTA 600
Qy      601 ACGTATTAACCGGAAATCTTTGAGAGTGAACATACCGGTTGACGTGGGAAACCGGACTC 660
Db      601 ACGTATTAACCGGAAATCTTTGAGAGTGAACATACCGGTTGACGTGGGAAACCGGACTC 660
Qy      661 GGACCGAAGTGAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
Db      661 GGACCGAAGTGAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
Qy      721 TTTTAAACAAAGTGTAGGCGGACGAAAGCCATTTGCTGCTGCGGATTTACGGAATATG 780
Db      721 TTTTAAACAAAGTGTAGGCGGACGAAAGCCATTTGCTGCTGCGGATTTACGGAATATG 780
Qy      781 GATATGTTGAGAGATGATTTCCCGGTTTGTGTGGGTTTATGCGCGGTTGATGATGATG 840
Db      781 GATATGTTGAGAGATGATTTCCCGGTTTGTGTGGGTTTATGCGCGGTTGATGATGATG 840
Qy      841 AATGGGCAAAAGAAAAAGTCTGAGATGAGAGCTTTGAGTGAATTTGTGCTTCCGTTT 900
Db      841 AATGGGCAAAAGAAAAAGTCTGAGATGAGAGCTTTGAGTGAATTTGTGCTTCCGTTT 900
Qy      901 CTCAAGTATGATTTGTGGGGTGAAGAAAGCGAGATTCATGATGATGATGATGATGATG 960
Db      901 CTCAAGTATGATTTGTGGGGTGAAGAAAGCGAGATTCATGATGATGATGATGATGATG 960
Qy      961 GTTTCTCCGGCCAGACTGATCTTCACTGATGATGATGATGATGATGATGATGATGATG 1020
Db      961 GTTTCTCCGGCCAGACTGATCTTCACTGATGATGATGATGATGATGATGATGATGATG 1020
Qy      1021 TAGATTTGTATGATCTATATTCAGAGGGGCTTGAATTTGAAAAACCTATCAATG 1080
Db      1021 TAGATTTGTATGATCTATATTCAGAGGGGCTTGAATTTGAAAAACCTATCAATG 1080
Qy      1081 TTTTCTAGCTCCAAAGCTAGCTATTTGTCATGCTTAAAGTTGATGATGATTTTATTTAA 1140
Db      1081 TTTTCTAGCTCCAAAGCTAGCTATTTGTCATGCTTAAAGTTGATGATGATTTTATTTAA 1140

```

```

Db      1081 TTTTCTAGCTCCAAAGCTAGCTATTTGTCATGCTTAAAGTTGATGATGATTTTATTTAA 1140
Qy      1141 CTGATCAAAACATTTGTATGATTTTAAACCAAAAAA 1188
Db      1141 CTGATCAAAACATTTGTATGATTTTAAACCAAAAAA 1188

RESULT 2
ADJ98180
ID ADJ98180 standard; DNA; 1188 BP.
XX
AC ADJ98180;
XX
DT 06-MAY-2004 (first entry)
XX
DE Thale cress chlorophyllase 1 DNA.
XX
KW phycol kinase; tocopherol biosynthesis; plant; drought resistance;
XX thale cress; ds; chlorophyllase 1.
XX
OS Arabidopsis thaliana.
XX
PN W020401312-A2.
XX
PD 12-FEB-2004.
XX
PR 05-AUG-2003; 2003MO-US025276.
XX
PR 05-AUG-2002; 2002US-040689P.
XX
PR 05-AUG-2003; 2003US-00634548.
XX
PA (MONS ) MONSANTO TECHNOLOGY LLC.
PI Norris SR, Lincoln K, Abad MS, Eilers R, Hartsuyker KK,
PI Hirschberg J, Karunandaa B, Moshiri F, Stein JC, Valentin HE,
PI Venkatesh TV;
XX
DR WPI; 2004-157125/15.
XX
PT New phycol kinase polynucleotides, useful in mediating tocopherol
PT biosynthesis and in producing plants with increased drought resistance.
XX
PS Claim 23; SEQ ID NO 18; 189pp; English.
XX
CC The invention relates to a novel substantially purified nucleic acid
CC molecule encoding a phycol kinase. The nucleic acid molecules and
CC polypeptides of the invention may be useful in mediating tocopherol
CC biosynthesis and in producing plants with increased drought resistance.
CC The current sequence is that of the thale cress chlorophyllase 1 DNA of
CC the invention.
XX
SQ Sequence 1188 BP; 327 A; 255 C; 281 G; 325 T; 0 U; 0 Other;

Query Match 100.0%; Score 1188; DB 12; Length 1188;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 GATCATTAATCTTCAACAACTCTTAATATCTACTTTAATACAAATGGCGCGATA 60
Db      1 GATCATTAATCTTCAACAACTCTTAATATCTACTTTAATACAAATGGCGCGATA 60
Qy      61 GAGGACAGTCCAGCGTTTCTCTGTGTAACTCCGCGCGCTTTGAGATAGCGACCTC 120
Db      61 GAGGACAGTCCAGCGTTTCTCTGTGTAACTCCGCGCGCTTTGAGATAGCGACCTC 120
Qy      121 CCGACAAACGAGATACCGGTGGATCCGGTGGAAAAATGATCAACGACCGCAAAAACG 180
Db      121 CCGACAAACGAGATACCGGTGGATCCGGTGGAAAAATGATCAACGACCGCAAAAACG 180
Qy      181 GTGAGAAATCACCTGTCCAAACAGTCGCGGAACTTAATCCCGTGTATTTCTTCATGCG 240
Db      181 GTGAGAAATCACCTGTCCAAACAGTCGCGGAACTTAATCCCGTGTATTTCTTCATGCG 240
Qy      241 TTTTATCTTGGCACTACTCTGATCTGACGTTCTTAACCAATCGCTTCCGATGCTTAC 300
Db      241 TTTTATCTTGGCACTACTCTGATCTGACGTTCTTAACCAATCGCTTCCGATGCTTAC 300
Qy      301 ATTCTGTAGCCCAAGTTGTGCAAAATTTAGCCGCGGAGAGGCAAGTGAAGTGGAC 360
Db      301 ATTCTGTAGCCCAAGTTGTGCAAAATTTAGCCGCGGAGAGGCAAGTGAAGTGGAC 360
Qy      361 GATGCTGAGAGTGTGATTAATCTGAGCATCGGAAAACTCAAAAGCTCACTCACTTAC 420
Db      361 GATGCTGAGAGTGTGATTAATCTGAGCATCGGAAAACTCAAAAGCTCACTCACTTAC 420
Qy      421 GTAAATGCTAATGAAAAATACACCTCACTCGTGGGCGCACGCGGGTGGAAAAACGGCG 480
Db      421 GTAAATGCTAATGAAAAATACACCTCACTCGTGGGCGCACGCGGGTGGAAAAACGGCG 480
Qy      481 TTGGCGTGTGGCGCTAGGCGCATGCGGCAACATTAGACCATCATCACTTTCAGCTCTA 540
Db      481 TTGGCGTGTGGCGCTAGGCGCATGCGGCAACATTAGACCATCATCACTTTCAGCTCTA 540
Qy      541 ATAGGAATTTGATCCAGTCGAGGAACTTAACAAATTCATTAGAACCGATCCGATATCTTA 600
Db      541 ATAGGAATTTGATCCAGTCGAGGAACTTAACAAATTCATTAGAACCGATCCGATATCTTA 600
Qy      601 ACGTATTAACCGGAAATCTTTGAGAGTGAACATACCGGTTGACGTGGGAAACCGGACTC 660
Db      601 ACGTATTAACCGGAAATCTTTGAGAGTGAACATACCGGTTGACGTGGGAAACCGGACTC 660
Qy      661 GGACCGAAGTGAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
Db      661 GGACCGAAGTGAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
Qy      721 TTTTAAACAAAGTGTAGGCGGACGAAAGCCATTTGCTGCTGCGGATTTACGGAATATG 780
Db      721 TTTTAAACAAAGTGTAGGCGGACGAAAGCCATTTGCTGCTGCGGATTTACGGAATATG 780
Qy      781 GATATGTTGAGAGATGATTTCCCGGTTTGTGTGGGTTTATGCGCGGTTGATGATGATG 840
Db      781 GATATGTTGAGAGATGATTTCCCGGTTTGTGTGGGTTTATGCGCGGTTGATGATGATG 840
Qy      841 AATGGGCAAAAGAAAAAGTCTGAGATGAGAGCTTTGAGTGAATTTGTGCTTCCGTTT 900
Db      841 AATGGGCAAAAGAAAAAGTCTGAGATGAGAGCTTTGAGTGAATTTGTGCTTCCGTTT 900
Qy      901 CTCAAGTATGATTTGTGGGGTGAAGAAAGCGAGATTCATGATGATGATGATGATGATG 960
Db      901 CTCAAGTATGATTTGTGGGGTGAAGAAAGCGAGATTCATGATGATGATGATGATGATG 960
Qy      961 GTTTCTCCGGCCAGACTGATCTTCACTGATGATGATGATGATGATGATGATGATGATG 1020
Db      961 GTTTCTCCGGCCAGACTGATCTTCACTGATGATGATGATGATGATGATGATGATGATG 1020
Qy      1021 TAGATTTGTATGATCTATATTCAGAGGGGCTTGAATTTGAAAAACCTATCAATG 1080
Db      1021 TAGATTTGTATGATCTATATTCAGAGGGGCTTGAATTTGAAAAACCTATCAATG 1080
Qy      1081 TTTTCTAGCTCCAAAGCTAGCTATTTGTCATGCTTAAAGTTGATGATGATTTTATTTAA 1140
Db      1081 TTTTCTAGCTCCAAAGCTAGCTATTTGTCATGCTTAAAGTTGATGATGATTTTATTTAA 1140

```

QY 241 TTTTATCTTGGCACTACTTCTAGCGTTCTTAACCAATCGCTTCCGATGGTTAC 300
DB 241 TTTTATCTTGGCACTACTTCTAGCGTTCTTAACCAATCGCTTCCGATGGTTAC 300
QY 301 ATTCTTGTAGCCCGACAGTTGTGCAATTATTTGCGCGGGAGGGAAGTGAAGTGAAC 360
DB 301 ATTCTTGTAGCCCGACAGTTGTGCAATTATTTGCGCGGGAGGGAAGTGAAGTGAAC 360
QY 361 GATGCTGGAAGTGTGATTAACCTGGGCAATCGGAAACCTCAAGCTCACTACCACTTCG 420
DB 361 GATGCTGGAAGTGTGATTAACCTGGGCAATCGGAAACCTCAAGCTCACTACCACTTCG 420
QY 421 GTAAATGCTAATGAAAAATACCTCACTCGTGGGCCACAGCCCGGTGGGAAAAACGCG 480
DB 421 GTAAATGCTAATGAAAAATACCTCACTCGTGGGCCACAGCCCGGTGGGAAAAACGCG 480
QY 481 TTTTCCGTTGGCTAGGCGCATGCGGCAATTAACCCATCAGTTTTCAGCTCTA 540
DB 481 TTTTCCGTTGGCTAGGCGCATGCGGCAATTAACCCATCAGTTTTCAGCTCTA 540
QY 541 ATAGAAATGATCCAGTCGAGAACTAACAAATACATTAGAACCGATCCGATATCTTA 600
DB 541 ATAGAAATGATCCAGTCGAGAACTAACAAATACATTAGAACCGATCCGATATCTTA 600
QY 601 ACGTATTAACCGGAAATCTTTGAGCTGGAATACCGGTGAGTGGTGGAAACCGGACTC 660
DB 601 ACGTATTAACCGGAAATCTTTGAGCTGGAATACCGGTGAGTGGTGGAAACCGGACTC 660
QY 661 GGAACGAAAGTGAACAAGTATGCAATGCGGCAACGAGCTTAACCATGAGAG 720
DB 661 GGAACGAAAGTGAACAAGTATGCAATGCGGCAACGAGCTTAACCATGAGAG 720
QY 721 TTTTACAAGAGTGAAGCGACGAAAGCCCATTTGCTGGCGAATTACGACATATG 780
DB 721 TTTTACAAGAGTGAAGCGACGAAAGCCCATTTGCTGGCGAATTACGACATATG 780
QY 781 GATATGTTGACGATGATTTGCGCGGTTTGTGGGTTATGCGCGGTGATGTGTAAG 840
DB 781 GATATGTTGACGATGATTTGCGCGGTTTGTGGGTTATGCGCGGTGATGTGTAAG 840
QY 841 AATGGGCAAGAAAAAGTGTGATGAGAGCTTTGTAGGTGGAATTTGTGCGCTT 900
DB 841 AATGGGCAAGAAAAAGTGTGATGAGAGCTTTGTAGGTGGAATTTGTGCGCTT 900
QY 901 CTCAAGTATAGTTGTGGGGTGAAGAAAGCGGAGATTCGATTTGTGAAGGATCTTCC 960
DB 901 CTCAAGTATAGTTGTGGGGTGAAGAAAGCGGAGATTCGATTTGTGAAGGATCTTCC 960
QY 961 GTTTCCTCCGCGCAAGCTTGATCCTTCACTGAGTTGGAAGAAAGCTTCTGGTATCTTCTC 1020
DB 961 GTTTCCTCCGCGCAAGCTTGATCCTTCACTGAGTTGGAAGAAAGCTTCTGGTATCTTCTC 1020
QY 1021 TAGATTTGTGTATGTAATATATCAAGAGGGGTCTTGAATTTTGAAGAAAGCTATCATG 1080
DB 1021 TAGATTTGTGTATGTAATATATCAAGAGGGGTCTTGAATTTTGAAGAAAGCTATCATG 1080
QY 1081 TTTTCTAGCTCAAGCTAGCTAATTTGATCATGCTCAAGTGCATGTATTTTATTA 1140
DB 1081 TTTTCTAGCTCAAGCTAGCTAATTTGATCATGCTCAAGTGCATGTATTTTATTA 1140
QY 1141 CTGATGAAACATTTGTTATAGTTTAACCCCAAAAAA 1188
DB 1141 CTGATGAAACATTTGTTATAGTTTAACCCCAAAAAA 1188

DE Arabidopsis thaliana DNA fragment SEQ ID NO: 5284.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
OS Arabidopsis thaliana.
PN EP103405-A2.
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 16-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135533P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.

PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144336P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144684P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 03-AUG-1999; 99US-0146389P.
PR 04-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.

PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 93.9%; Score 115.4; DB 3; Length 1157;

Best Local Similarity 99.5%; Pred. No. 0; Mismatches 1; Indels 5; Gaps 3;

Matches 1151; Conservative 0; Mismatches 1; Indels 5; Gaps 3;

QY 2 ATACATAATCTTCAACAACAACCTTTAATTATCTAGTTTAAATCAAAATGGCGGCGATAG 61
DB 5 ATACATAATCTTCAACAACAACCTTTAATTATCTAGTTTAAATCAAAATGGCGGCGATAG 64
QY 62 AGGACATCCAACTTTTCTCTGTGTGTAATCCGGGGGCTTTTGAATATGGCAAGCTCC 121
DB 65 AGGACATCCAACTTTTCTCTGTGTGTAATCCGGGGGCTTTTGAATATGGCAAGCTCC 124
QY 122 CGACAAACGAGATACCGGTGATCCGGTGGAAATGATTCACACGACCGCAAAACCGG 181
DB 125 CGACAAACGAGATACCGGTGATCCGGTGGAAATGATTCACACGACCGCAAAACCGG 184
QY 182 TGAAGATCACCTGTCAACAGTCGCGGAACTTATCCGCTGTTTATTTTCATGAGCT 241
DB 185 TGAAGATCACCTGTCAACAGTCGCGGAACTTATCCGCTGTTTATTTTCATGAGCT 242
QY 242 TTTATCTTGGCAATCTTCTGACTGTGACGTTCTTAACACATGCTTTCGATGTTTCA 301
DB 243 TTTATCTTGGCAATCTTCTGACTGTGACGTTCTTAACACATGCTTTCGATGTTTCA 302

```

Oy 302 TTCTGTAGCCCAAGTGTGCAATTATTTGCCCGCGGAGGCAATGGAATGGAAGCG 361
Db 303 TTTCTGTAGCCC--CAGTTGTGCAAAATTAATGCGCGCGGAGGCAATGGAATGGAAGCG 360
Oy 362 ATGCTGGAAGTGTGATTAACCTGAGCATCGAAAACTGAAAGCTCACTCAACACTTGG 421
Db 361 ATGCTGGAAGTGTGATTAACCTGAGCATCGAAAACTGAAAGCTCACTCAACACTTGG 420
Oy 422 TAAATGCTAATGAAAAATACACTCACTGCTGGGCGCAACGCGCGGTGGAAAAAGCGCGT 481
Db 421 TAAATGCTAATGAAAAATACACTCACTGCTGGGCGCAACGCGCGGTGGAAAAAGCGCGT 480
Oy 482 TTGGGGTGGCGCTAGGCGCATGCGCAATTAAGACCATTAACAGCTTTTCAAGTCTTAA 541
Db 481 TTGGGGTGGCGCTAGGCGCATGCGCAATTAAGACCATTAACAGCTTTTCAAGTCTTAA 540
Oy 542 TAGGAATTTGATCCAGTCCGAGAACTTAACAATAATACATTAAGAACGATCCGATATCTTAA 601
Db 541 TAGGAATTTGATCCAGTCCGAGAACTTAACAATAATACATTAAGAACGATCCGATATCTTAA 600
Oy 602 CGTATTAACCGGGAATCTTTTCAGCTGGAATACCGGTTTGAAGTGGGAAACCGGACTCG 661
Db 601 CGTATTAACCGGGAATCTTTTCAGCTGGAATACCGGTTTGAAGTGGGAAACCGGACTCG 660
Oy 662 GACCGAAGTGAACAAGCTGATGCCATGCGACCAACGGAATTTAAACATGAGAGT 721
Db 661 GACCGAAGTGAACAAGCTGATGCCATGCGACCAACGGAATTTAAACATGAGAGT 720
Oy 722 TTTACAAAGAGTGAAGCGAGCAAGAAAGCCATTTGCTGGCTGCGGATTAACGGAATATGG 781
Db 721 TTTACAAAGAGTGAAGCGAGCAAGAAAGCCATTTGCTGGCTGCGGATTAACGGAATATGG 780
Oy 782 ATATGTTGAGCAGATGATTTGCCCGGTTTGTGGGTTTATGCGCGGTTGATGTTGAAGA 841
Db 781 ATATGTTGAGCAGATGATTTGCCCGGTTTGTGGGTTTATGCGCGGTTGATGTTGAAGA 839
Oy 842 ATGGGCAAGAAAAAGCTGAGATGAGAGCCTTTGATGAGTGAATTTGCTGGCTTTC 901
Db 840 ATGGGCAAGAAAAAGCTGAGATGAGAGCCTTTGATGAGTGAATTTGCTGGCTTTC 899
Oy 902 TCAAGTATGATTTGTGGGCTGAAAAAGCGAGATTTGATGATTTGTAAGGATCTTCCG 961
Db 900 TCAAGTATGATTTGTGGGCTGAAAAAGCGAGATTTGATGATTTGTAAGGATCTTCCG 959
Oy 962 TTTCTCCGCGCAAGCTGATCTTCACTGAGTTGGAAGAACTTCTGATATCTTGGCT 1021
Db 960 TTTCTCCGCGCAAGCTGATCTTCACTGAGTTGGAAGAACTTCTGATATCTTGGCT 1019
Oy 1022 AGATTTGATGATGATGATGATGATGAGGCGCTTGAATTTGAAAAAAGCTATCAATGT 1081
Db 1020 AGATTTGATGATGATGATGATGATGAGGCGCTTGAATTTGAAAAAAGCTATCAATGT 1079
Oy 1082 TTTCTAGCTCCAGACTAGTATTTGATGATGATGATGATGATGATGATGATGATGATGAT 1141
Db 1080 TTTCTAGCTCCAGACTAGTATTTGATGATGATGATGATGATGATGATGATGATGATGAT 1139
Oy 1142 TCGATCAAAACATTTGT 1158
Db 1140 TCGATCAAAACATTTGT 1156

```

RESULT 4
AB212353 standard; DNA; 975 BP.

AB212353;

21-JAN-2003 (first entry)

Arabidopsis thaliana stress regulated gene SEQ ID NO 158.

Arabidopsis thaliana; plant; gene; stress; transgenic; de.

XX

OS Arabidopsis thaliana.

XX WO200216655-A2.

XX 28-FEB-2002.

XX 24-AUG-2001; 2001WO-US026685.

XX 24-AUG-2000; 2000US-0227866P.

XX 26-JAN-2001; 2001US-0264647P.

XX 22-JUN-2001; 2001US-0300111P.

XX (Scrip) SCRIPPS RES INST.

XX (SYGN) SYNGENTA PARTICIPATIONS AG.

XX Harper JF, Kreps J, Wang X, Zhu T;

XX WPI; 2002-304127/34.

XX Claim 144; SEQ ID NO 158; 577bp + Sequence Listing; English.

XX The invention relates to identifying a stress condition to which a plant

cell has been exposed, comprising: (a) contacting nucleic acid

representative of expressed polynucleotides in the plant cell with an

array or probes representative of the plant cell genome; and (b)

detecting a profile of expressed polynucleotides in the plant cell

characteristic of a stress response. The method is useful in the

production of transgenic plants, cells and seeds and in producing plants

with increased tolerance to abiotic stress. The present sequence is that

of an Arabidopsis thaliana stress regulated gene (AB212196-AB217574) used

in methods of the invention. Note: The sequence data for this patent is

not represented in the printed specification but is based on sequence

information supplied to Derwent by the European Patent Office

XX Sequence 975 BP; 250 A; 222 C; 256 G; 247 T; 0 U; 0 Other;

XX Query Match 82.1%; Score 975; DB 6; Length 975;

XX Best Local Similarity 100.0%; Pred. No. 3.9e-285;

XX Matches 975; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 49 ATGGCGCGATGAGAGCAGTCCAGCTTTTCTGTGTGTAATCCGCGGCTTTTGAAG 108

Db 1 ATGGCGCGATGAGAGCAGTCCAGCTTTTCTGTGTGTAATCCGCGGCTTTTGAAG 60

Oy 109 ATAGGCGAGCTCCCGCAACCGAGATACCGGTGATCCGCTGAAAAATGATCAACGCA 168

Db 61 ATAGGCGAGCTCCCGCAACCGAGATACCGGTGATCCGCTGAAAAATGATCAACGCA 120

Oy 169 CCGCCAAAACCGGTGAGAAATCACTGTCACAGTGCAGGAACTTATCCCGTGTTTTA 228

Db 121 CCGCCAAAACCGGTGAGAAATCACTGTCACAGTGCAGGAACTTATCCCGTGTTTTA 180

Oy 229 TTTCTCCAGCTTTTATCTTTCGCAACTACTTCTGACGTTCTTAACCAATCGCT 288

Db 181 TTTCTCCAGCTTTTATCTTTCGCAACTACTTCTGACGTTCTTAACCAATCGCT 240

Oy 289 TCGCATGTTACATTTCTGTAGGCCCAAGTTGTGAAAAATATTTGCGCGGAGAGGCA 348

Db 241 TCGCATGTTACATTTCTGTAGGCCCAAGTTGTGAAAAATATTTGCGCGGAGAGGCA 300

Oy 349 GTGGAAGTGAACGATCGTGAAGTGAATTAACCTGGGCAATCGAAAAAGCTCAAGCTCAC 408

Db 301 GTGGAAGTGAACGATCGTGAAGTGAATTAACCTGGGCAATCGAAAAAGCTCAAGCTCAC 360

Oy 409 CTACCACTTGGTAAATGCTAATGAAAAATACCTCACTGCTGGGCAACGCGCGGT 468

Db 361 CTACCACTTGGTAAATGCTAATGAAAAATACCTCACTGCTGGGCAACGCGCGGT 420

Oy 469 GGAAGAAACGCGCTTTCGCGGTGCTAGGCGCATGCGCAACATTAACCATCATCAG 528

Db 469 GGAAGAAACGCGCTTTCGCGGTGCTAGGCGCATGCGCAACATTAACCATCATCAG 528

```
Db 421 GGGAAACGGCGTTTGCGGTTGCGTAGGCCATGCCGACATTAGACCATTCCATCCG 480
Qy 529 TTTTACGCTCTAATGGAATGATCCAGTCGACGAACTAACAAATACATTAGAACGAT 588
Db 481 TTTTACGCTCTAATGGAATGATCCAGTCGACGAACTAACAAATACATTAGAACGAT 540
Qy 589 CCGCATCTTAAACGATTAACCCGGAATCTTTCCAGCTGACATCCGCTGGCAGTGGTG 648
Db 541 CCGCATCTTAAACGATTAACCCGGAATCTTTCCAGCTGACATCCGCTGGCAGTGGTG 600
Qy 649 GGAACCGGACTCGACCGAAGTGAACAACGATGACCAATGCGACCAACGACTTA 708
Db 601 GGAACCGGACTCGACCGAAGTGAACAACGATGACCAATGCGACCAACGACTTA 660
Qy 709 AACCATGAGAGTTTACAAAGAGTGAAGGAGGAGAAAGCCCATTTGTGGCTGCGAT 768
Db 661 AACCATGAGAGTTTACAAAGAGTGAAGGAGGAGAAAGCCCATTTGTGGCTGCGAT 720
Qy 769 TACGACATATGATATGATGAGACGATATTTGCCGGTTTGTGGGTTTATGGCCGAT 828
Db 721 TACGACATATGATATGATGAGACGATATTTGCCGGTTTGTGGGTTTATGGCCGAT 780
Qy 829 TGTATGTATGAAGATGGGCAAAAGAAAAAGTCTGAGATGAGAGCTTTGTAGTGAATT 888
Db 781 TGTATGTATGAAGATGGGCAAAAGAAAAAGTCTGAGATGAGAGCTTTGTAGTGAATT 840
Qy 889 GTGCTGTGCTTTCTCAAGTATATGTTGTGGGTTGAAAAAGCGGAGATTCGATTGTTG 948
Db 841 GTGCTGTGCTTTCTCAAGTATATGTTGTGGGTTGAAAAAGCGGAGATTCGATTGTTG 900
Qy 949 AAGGATCTCTCCGTTTCCGGCCAAAGCTGATCCTTCACTGAGTTGGAAGAGCTTCT 1008
Db 901 AAGGATCTCTCCGTTTCCGGCCAAAGCTGATCCTTCACTGAGTTGGAAGAGCTTCT 960
Qy 1009 GGTATCTTCTGCTAG 1023
Db 961 GGTATCTTCTGCTAG 975

RESULT 5
ADG87791
ID ADG87791 standard; cDNA; 975 BP.
XX
AC ADG87791;
XX
DT 22-APR-2004 (first entry)
XX
DE A. thaliana RPP7/RPP8-upregulated pathogen infection-related gene #233.
XX
KW Pathogen infection-related gene; plant; Peronospora parasitica;
KW defence mechanism; RPP7; RPP8; pathogen resistance; transgenic plant;
KW oomycete; fungus; bacterium; virus; nematode; insect; aphid; gene; ss.
XX
OS Arabidopsis thaliana.
XX
PN MO200222675-AZ.
XX
PD 21-MAR-2002.
XX
PE 14-SEP-2001; 2001WO-US028506.
XX
PR 15-SEP-2000; 2000US-0232778P.
XX
PR 22-JUN-2001; 2001US-0300183P.
XX
PA (SYN) SYNGENTA PARTICIPATIONS AG.
PA (UNCL) UNIV NORTH CAROLINA.
PA (GLAZ) GLAZEBROOK J.
PA (WANG) WANG X.
PA (DANG) DANG J L.
PA (EULG) EULGEN T.
PA (ZHUT) ZHU T.
PI Glazebrook J, Wang X, Dang J L, Eulgen T, Zhu T;
```

```
XX
XX WPI: 2002-292409/33.
XX
XX Novel isolated polynucleotide, useful for conveying pathogen resistance
XX to plants, and for identifying plants infected with a pathogen.
XX
XX Claim 3; SEQ ID NO 233; 605bp; English.
XX
XX The invention relates to 691 Arabidopsis thaliana genes (ADG87559--
XX ADG87557)) whose expression is altered in response to pathogen infection,
XX or to homologues of these genes from other plants or fungi, especially
XX from maize, soybean, barley, alfalfa, sunflower, canola (oilseed rape),
XX cotton, peanut, sorghum, tobacco, sugarbeet, rice or wheat. The
XX expression of genes of the invention was upregulated or downregulated in
XX Arabidopsis plants infected with the oomycete Peronospora parasitica,
XX indicating that they play a role in defence mechanisms. The genes of the
XX invention are regulated by RPP7 or RPP8 which act via unconventional
XX signalling cascades, or by the RPP4-dependent pathway. The invention also
XX relates to polypeptides encoded by the pathogen infection-related genes;
XX promoter motifs from pathogen infection-related genes (ADG88243-ADG88327)
XX; expression cassettes, host cells and pathogen-resistant transgenic
XX plants and their progeny comprising a polynucleotide of the invention;
XX and a method of identifying a plant cell infected with a pathogen. The
XX polynucleotide sequences and methods of the invention are useful for
XX identifying plants infected with a pathogen, and for conferring
XX resistance to pathogens such as oomycetes, fungi, bacteria, viruses,
XX nematodes and insects (e.g., aphids). The present sequence represents an
XX Arabidopsis thaliana gene whose expression is altered in response to
XX Peronospora parasitica infection. Note: The sequence data for this patent
XX did not form part of the printed specification, but was obtained in
XX electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 975 BP; 250 A; 222 C; 256 G; 247 T; 0 U; 0 Other;
XX
Query Match 82.1%; Score 975; DB 6; Length 975;
Best Local Similarity 100.0%; Pred. No. 3.9e-285;
Matches 975; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 49 ATGGCGGCGATAGAGGACAGTCCAAAGTTTCTCTGTGTAATCCGGCGGCTTTTGAG 108
Db 1 ATGGCGGCGATAGAGGACAGTCCAAAGTTTCTCTGTGTAATCCGGCGGCTTTTGAG 60
Qy 109 ATAGGACAGCTCCCGAACAACGAGATACCGGTGATCCGGTGAATAATGATTCAACAGCA 168
Db 61 ATAGGACAGCTCCCGAACAACGAGATACCGGTGATCCGGTGAATAATGATTCAACAGCA 120
Qy 169 CCGCAAAACCGGTGAGATACCTGTCAACAGTCCCGGAATTATCCCGTGTTTA 228
Db 121 CCGCAAAACCGGTGAGATACCTGTCAACAGTCCCGGAATTATCCCGTGTTTA 180
Qy 229 TTTCTTCATGGCTTTTATCTTCGCACTCTTCACTGTGACGTTCTTAACCAATGGCT 288
Db 181 TTTCTTCATGGCTTTTATCTTCGCACTCTTCACTGTGACGTTCTTAACCAATGGCT 240
Qy 289 TCGCATGTTAATCTTTGTAGCCCAAGTTGTCAAAATTAATTCGCGGAGGGCAA 348
Db 241 TCGCATGTTAATCTTTGTAGCCCAAGTTGTCAAAATTAATTCGCGGAGGGCAA 300
Qy 349 GTGGAAGTGACGATGCTGGAAGTGTATAACTGGGCAATCGAANAACCTCAAGCTTAC 408
Db 301 GTGGAAGTGACGATGCTGGAAGTGTATAACTGGGCAATCGAANAACCTCAAGCTTAC 360
Qy 409 CTACCAACTTCGTAATGCTAAATGAAATACACTCTCGTGGGCGACACGCGGCT 468
Db 361 CTACCAACTTCGTAATGCTAAATGAAATACACTCTCGTGGGCGACACGCGGCT 420
Qy 469 GGAAGAAACGGGCTTTGGGCTGAGCCATCGCAACATTAGACCCATCCATCAGC 528
Db 421 GGAAGAAACGGGCTTTGGGCTGAGCCATCGCAACATTAGACCCATCCATCAGC 480
Qy 529 TTTTCACTCTAATGAAATGATCCAGTCGACAGAACTAACAAATACATTAGAACGAT 588
Db 589 TTTTCACTCTAATGAAATGATCCAGTCGACAGAACTAACAAATACATTAGAACGAT 540
```

Db 481 TTTTCAGCTTAATAGAAATTGATCCAGTCCGACGAACTAACAATACATTAGAACCGAT 540
Qy 589 CCGCATATCTTAAAGCTATTAACCGGAATCTTTCCAGCTGACATACCGGTTGACAGTGTG 648
Db 541 CCGCATATCTTAAAGCTATTAACCGGAATCTTTCCAGCTGACATACCGGTTGACAGTGTG 600
Qy 649 GGAACCCGAGCTCGGACCGGAAGTGAACAACGTGATGCCACATCCGACCAACGACACTTA 708
Db 601 GGAACCCGAGCTCGGACCGGAAGTGAACAACGTGATGCCACATCCGACCAACGACACTTA 660
Qy 709 AACCATGAGAGCTTTTACAAAGAGTGAAGGCGACGAAGGCCAATTTGCTGGCTGCGGAT 768
Db 661 AACCATGAGAGCTTTTACAAAGAGTGAAGGCGACGAAGGCCAATTTGCTGGCTGCGGAT 720
Qy 769 TACGACATATGATGATGTTGACAGCATATTTGCCGGTTTGTGGGTTTATGACCGGT 828
Db 721 TACGACATATGATGATGTTGACAGCATATTTGCCGGTTTGTGGGTTTATGACCGGT 780
Qy 829 TGTATGTATAGAAATGGGCAAGAAAGAGTGAAGTGAAGGACCTTTGTAGTGGGAAT 888
Db 781 TGTATGTATAGAAATGGGCAAGAAAGAGTGAAGTGAAGGACCTTTGTAGTGGGAAT 840
Qy 889 GTGGTTGGCTTCTCAAGTATAGTTTGTGGGTTGAAAAAGCGGAGATTGATGATTGTG 948
Db 841 GTGGTTGGCTTCTCAAGTATAGTTTGTGGGTTGAAAAAGCGGAGATTGATGATTGTG 900
Qy 949 AAGATCTCTTCGTTTCTCCGCGCAAGCTTGATCTTCACTGATGGAAGAAAGCTTCT 1008
Db 901 AAGATCTCTTCGTTTCTCCGCGCAAGCTTGATCTTCACTGATGGAAGAAAGCTTCT 960
Qy 1009 GGATCTTCGTTCTAG 1023
Db 961 GGATCTTCGTTCTAG 975

RESULT 6
ADG87590
ID ADG87590 standard; cDNA; 975 BP.
XX
AC ADG87590;
XX
DT 22-APR-2004 (first entry)
XX
DE A. thaliana RPP7/RPP8-upregulated pathogen infection-related gene #32.
XX
KW Pathogen infection-related gene; plant; Peronospora parasitica;
KW defence mechanism; RPP7; RPP8; pathogen resistance; transgenic plant;
KW oomycete; fungus; bacterium; virus; nematode; insect; aphid; gene; ss.
OS Arabidopsis thaliana.
XX
PN WC200222675-A2.
XX
PD 21-MAR-2002.
XX
PE 14-SEP-2001; 2001MO-US028506.
XX
PR 15-SEP-2000; 2000US-0232778P.
XX
PR 22-JUN-2001; 2001US-0300183P.
XX
PA (SYGN) SYNGENTA PARTICIPATIONS AG.
PA (UYNC-) UNIV NORTH CAROLINA.
PA (GLAZ/) GLAZEBROOK J.
PA (WANG/) WANG X.
PA (DANG/) DANG J L.
PA (EULGEM T. EULGEM T.
PA (ZHU/) ZHU T.
XX
PI Glazebrook J, Wang X, Dang J L, Bulgem T, Zhu T,
XX
XX WPI; 2002-292409/33.
XX
PT Novel isolated polynucleotide, useful for conveying pathogen resistance

PT to plants, and for identifying plants infected with a pathogen.
XX
XX Claim 3; SEQ ID NO 32; 605bp; English.
PS
XX
XX The invention relates to 691 Arabidopsis thaliana genes (ADG87559--
CC ADG87557) whose expression is altered in response to pathogen infection,
CC and to homologues of these genes from other plants or fungi, especially
CC from maize, soybean, barley, alfalfa, sunflower, canola (oilseed rape),
CC cotton, peanut, sorghum, tobacco, sugarcane, rice or wheat. The
CC expression of genes of the invention was upregulated or downregulated in
CC Arabidopsis plants infected with the oomycete Peronospora parasitica.
CC Arabidopsis plants infected with the complete Peronospora parasitica,
CC indicating that they play a role in defence mechanisms. The genes of the
CC invention are regulated by RPP7 or RPP8 which act via unconventional
CC signalling cascades, or by the RPP4-dependent pathway. The invention also
CC relates to polypeptides encoded by the pathogen infection-related genes;
CC promoter motifs from pathogen infection-related genes (ADG88243-ADG88327)
CC ; expression cassettes, host cells and pathogen-resistant transgenic
CC plants and their progeny comprising a polynucleotide of the invention;
CC and a method of identifying a plant cell infected with a pathogen. The
CC polynucleotide sequences and methods of the invention are useful for
CC identifying plants infected with a pathogen, and for conferring
CC resistance to pathogens such as oomycetes, fungi, bacteria, viruses,
CC nematodes and insects (e.g., aphids). The present sequence represents an
CC Arabidopsis thaliana gene whose expression is altered in response to
CC Peronospora parasitica infection. Note: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
SQ Sequence 975 BP; 250 A; 222 C; 256 G; 247 T; 0 U; 0 Other;
Query Match 82.1%; Score 975; DB 6; Length 975;
Best Local Similarity 100.0%; Pred. No. 3.9e-285;
Matches 975; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 49 ATGGCGGCGATATAGAGGACAGTCCAGCTTTCTCTGTGTACTCCGGCGGCTTTTGTAG 108
Db 1 ATGGCGGCGATATAGAGGACAGTCCAGCTTTCTCTGTGTACTCCGGCGGCTTTTGTAG 60
Qy 109 ATAGGAGCCTCCCGCAACCGAGATACCGGTGATCCGTGGAATAATGTTCAACGACA 168
Db 61 ATAGGAGCCTCCCGCAACCGAGATACCGGTGATCCGTGGAATAATGTTCAACGACA 120
Qy 169 CCGCCAAACCGGTGAGATACCTGTCCACAGTGGCGGAACCTATCCCGTGTTTTA 228
Db 121 CCGCCAAACCGGTGAGATACCTGTCCACAGTGGCGGAACCTATCCCGTGTTTTA 180
Qy 229 TTCTTCATGAGCTTTTATCTTGGCACTACTTCTACTGAGCTGTTAACCAATCCGCT 288
Db 181 TTCTTCATGAGCTTTTATCTTGGCACTACTTCTACTGAGCTGTTAACCAATCCGCT 240
Qy 289 TCGCATGTTTACATTTCTTGTAGCCCAAGTTGTGAATAATTATGCGCGCGGAGGCA 348
Db 241 TCGCATGTTTACATTTCTTGTAGCCCAAGTTGTGAATAATTATGCGCGCGGAGGCA 300
Qy 349 GTGGAAGTGAAGCATGCTGGAAGTGTATTAACCTGAGGAAACCTCAAAAGCTCAC 408
Db 301 GTGGAAGTGAAGCATGCTGGAAGTGTATTAACCTGAGGAAACCTCAAAAGCTCAC 360
Qy 409 CTACCAACTTCGTTAATGCTAATGGAATAATACCTCACTCGTGGGGCCACACCGCGGT 468
Db 361 CTACCAACTTCGTTAATGCTAATGGAATAATACCTCACTCGTGGGGCCACACCGCGGT 420
Qy 469 GGGAAAAACGGCGTTTGGCGTTGGCGTAGGCGCATGCGCAACATTAGACCATCATCAG 528
Db 421 GGGAAAAACGGCGTTTGGCGTTGGCGTAGGCGCATGCGCAACATTAGACCATCATCAG 480
Qy 529 TTTTCAGCTTAAATAGAAATTGATCAGTGGCAGGAACCTTAACAATATAGAACCGAT 588
Db 481 TTTTCAGCTTAAATAGAAATTGATCAGTGGCAGGAACCTTAACAATATAGAACCGAT 540
Qy 589 CCGCATATCTTAAAGCTATTAACCGGAATCTTTCCAGCTGACATACCGGTTGACAGTGTG 648

Dh 611 CGACCTGAGCATGCCGGTTGAGTATGCTGATCGGCTCTCGAGCCGGAAGTAACATGCT 670
Qy 661 GATGCCACCATTCGCAACCAACGACTTAAACCATAGAGATTTTACAAAGGTGAAGGC 740
Db 671 GATGCCACCATTCGCAACCAACGACTTAAACCATAGAGATTTTATTTAGGTGAAGGC 730
Qy 741 GACGAAAGCCCATTTGCTGGCTGGGATTAAGGACATATGATATGTTGAGCGATGATT 800
Db 731 TACCAAGGGGCAATTTGCTGGCTGGGATTAAGGACATATGATATGTTGAGCGATATTT 790
Qy 801 GCCCGGTTTGTGGGTTTATGCGCGGTGTATGTGTATGTAAGATGGCAAAAGAAAGTTC 860
Db 721 GCCCGGTTTGTGGGTTTATGCGCGGTGTATGTGTATGTAAGATGGCAAAAGAAAGTTC 850
Qy 861 TGAGATGAGAGCTTTTGAAGTGTGAATGTGTTGCTTCAAGTATGTTGTGGGG 920
Db 851 TGAGATGAGAGCTTTTGAAGTGTGAATGTGTTGCTTCAAGTATGTTGTGGGG 909
Qy 921 TGAAAAAGCGGAGATTGATGATGTAAGGATGCTCCGTTTCTCGGCGCAAGCTTGA 980
Db 910 TGAAAAAGCGGAGATTGATGATGTAAGGATGCTCCGTTTCTCGGCGCAAGCTTGA 969
Qy 981 TCCTTACCTGATGTGAAGAGCTTCTGTATCTTCTGCTAGATTTGTGTATGTAAGT 1040
Db 970 TCCTTACCTGATGTGAAGAGCTTCTGTATCTTCTGCTAGATTTGTGTATGTAAGT 1029
Qy 1041 TTATCAAGGGGCTCTTGAATTTTGAAGAACTTATCAATGTTT 1084
Db 1030 GTATTGTATCTCTCAGGAGAGCTCCAACTGTAATAAATTAT 1073

RESULT 8

AAD38074
ID AAD38074 standard; cDNA; 987 BP.

XX AAD38074;

AC 10-SEP-2002 (first entry)

DE Grape chlorophyllase cDNA from vdbic.pk002.p19:fls clone.

KW Grape; chlorophyllase; chlorophyll degradation; plant cell senescence;
enzyme; gene; ss.

XX Vitis sp.

OS Location/Qualifiers

XX Key

XX CDS

1..858
/product= "Grape chlorophyllase protein from
vdbic.pk002.p19:fls clone"
/EC_number= "3.1.1.14"
/note= "CDS does not include start codon"
/partial

XX MO200229022-A2.

XX 11-APR-2002.

XX 04-OCT-2001; 2001WO-US031059.

XX 05-OCT-2000; 2000US-0238161P.

XX (DUPO) DU PONT DE NEMOURS & CO E. I.

XX Cahoon EB, Cahoon RE, Thorpe CJ;

XX WPI; 2002-444102/47.

XX P-PSDB; AAE23776.

XX An isolated polynucleotide encoding a plant chlorophyllase polypeptide,
PT used to produce transformed plants that have controlled induction or
PT postponement of senescence.

XX Claim 5; Page 47; 69pp; English.
PS The invention relates to plant chlorophyllase (EC 3.1.1.14) proteins and
XX polynucleotides encoding such proteins. chlorophyllase enzyme is involved
CC in chlorophyll degradation in plant cell senescence. Sequences of the
CC invention are used to produce a plant by transforming a plant cell with
CC chlorophyllase and regenerating a plant from the transformed plant cell.
CC They may also be used to transform cells. The plants that are produced
CC can have economical importance as they can allow for the controlled
CC induction or postponement of senescence. The present sequence is grape
CC chlorophyllase cDNA from vdbic.pk002.p19:fls clone
XX
SQ Sequence 987 BP; 260 A; 224 C; 218 G; 285 T; 0 U; 0 Other;
Query Match 20.2%; Score 239.4; DB 6; Length 987;
Best Local Similarity 57.6%; Pred. No. 1.3e-61;
Matches 473; Conservative 0; Mismatches 336; Indels 12; Gaps 2;
Qy 167 CACGCCAAAACCGGTGAGATCACTGTCCAAAGTCCGCGAATTATCCGTCGTTT 226
Db 26 CCCCCTCCCAAGCCATTGTGATGTTTACACCAACATTCAAGGAGACATACCAAGTCTCT 85
Qy 227 TATTCCTCAAGGCTTTTATCTTCCGCAACTTCTAAGTCTGACGTTCTTAACCAATCG 286
Db 86 TGTTCCTCAAGGCTTTTATCTTCCGCAACTTCTAAGTCTGACGTTCTTAACCAATCG 145
Qy 287 CTTGCGAATGATCATTTCTGTATGAGCCCCACAGTGTGCAATTATGCGCGGAGAGGC 346
Db 146 CTTCCATGATTCATTTGTGTGTGCTCTCAAGTATACGACATATTAATCTCTTGTGAA 205
Qy 347 AAGTGAAGTGAAGATGCTGGAAGTGTATTAATCTGCGATGGAATCTCAAGCTC 406
Db 206 TTCAAGGATCAATCAAGCAGCAGCACTCAAAATGCTATCTCAGGCTTCAATCTG 265
Qy 407 ACCTACCACTCGGTAAATGCTAATGGAATATACCTCACTCGTGGGCGACAGCGCG 466
Db 266 TGCTCCAGAAATGTGAACCAAGACCTTCAAGCTTGTCTTTACGGCGACAGCAGAG 325
Qy 467 GTGGAAACCGGCTTTCGCTGCGTACGCGATCCGCAAGTATGAGCCATTCATCA 526
Db 326 GGGGAAAGACAGCATTTGCTGTGCACTAGAGATG-----TGATACATCCCTCA 376
Qy 527 CGTTTACGCTTATATGGAATGATGATCAGTCCGAGAACTTAACAAATACATTAAGACG 586
Db 377 ACTTCTCAGCCCTACTAGAGACTAGACCTCTGTGTGAGTGAATATGTTGCCAAACAG 436
Qy 587 ATCCGATATCTTAAGGTATTAACCGGAATCTTTCAGCTGAGCATACCGGTGAGTGG 646
Db 437 TTCCCAAAATCTTAACCTATGTTCTCTCATCTTCAATCTAGCAATCCAGTTTGGCTAA 496
Qy 647 TGGGAACCGGACTCGACCGAG---TGAACAACAGTATGCCACATGCGCAACAGG 703
Db 497 TCGGCAACGGGTTTGGGAGATGAGCAAGAACTGCTTAATATGATGATGATGATG 556
Qy 704 ACTTAAACATGAGAGTTTACAAAGAGTGAAGCGACGAAAGCCCATTTGTGCTG 763
Db 557 GAGTGAACCATGTGAGATTTCAGTGAAGTGAACCTCTGTCTCCACTTTGTGACTA 616
Qy 764 CGGATTAACGACATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 823
Db 617 CTGAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 676
Qy 824 CCGGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 883
Db 677 CGGTTATATCTCAAGAGTGAAGAGGCTCTAGGAGCCCATGAGAGATGATGATGATG 736
Qy 884 GAATTTGCTGCTTTCTCAAGTATGATTTGTGGGTTGAAAAAGCGAGATTCATGA 943
Db 737 GCCTTTTGTGATTTCTGAAGGCTTATTTGGAAGTCAAGCTGAGATTTCAAAAGCCA 796
Qy 944 TTGGAAGATCTCTCGTTTCCGCGCAAGCTTATGCT 984


```

XX 11-APR-2002.
PD 04-OCT-2001; 2001WO-US031059.
PF 05-OCT-2000; 2000US-0238161P.
PR (DUPO ) DU PONT DE NEMOURS & CO E I.
XX Cahoon EB, Cahoon RE, Thorpe CJ;
XX WPI; 2002-444102/47.
DR P-PSDB; AAE23784.
XX An isolated polynucleotide encoding a plant chlorophyllase polypeptide,
PT used to produce transformed plants that have controlled induction or
PS postonement of senescence.
PS Claim 5; Page 59; 69pp; English.
XX The invention relates to plant chlorophyllase (EC 3.1.1.14) proteins and
CC polynucleotides encoding such proteins. Chlorophyllase enzyme is involved
CC in chlorophyll degradation in plant cell senescence. Sequences of the
CC invention are used to produce a plant by transforming a plant cell with
CC chlorophyllase and regenerating a plant from the transformed plant cell.
CC They may also be used to transform cells. The plants that are produced
CC can have economical importance as they can allow for the controlled
CC induction or postonement of senescence. The present sequence is wheat
CC chlorophyllase cDNA from wleln.pk0058.a4:fls clone
XX
SQ Sequence 1242 BP; 266 A; 382 C; 343 G; 251 T; 0 U; 0 Other;
Query Match 16.8%; Score 199; DB 6; Length 1242;
Beet Local Similarity 55.8%; Pred. No. 2.8e-49;
Matches 476; Conservative 0; Mismatches 350; Indels 27; Gaps 4;
QY 137 CGGTGATCCGGTGGAAATGATTCACAGACCGCCAAACCGGTGAGATCCTGTC 196
DB 153 CGATTCAAGTGAATGAGATCGGGCCGACACCAACGATCCGGTGTGATCGTCAC 212
QY 197 CAACAGTCGCCGGAATCTTCCGTTTATCTTCATGAGGCTTTATCTTCGCAACT 256
DB 213 CCAAGATGACGAACTACCCCGTGGCCATGCTTTCAGAGGCTTCTCCATTAAC 272
QY 257 ACTTCACTGACGTTCTTAACAACATCGCTTGCATGTTACATTTGTAGCCCCAC 316
DB 273 ACTTCACTGAAACCTTCTCGGACGTCGATCCACGCGCTTCATATGTCCGCCCC 332
QY 317 AGTTGTG---CAATTAATTGCGCGCGGAGGCGAATGGAAGTGAAGTCTGAAAGTG 373
DB 333 AGTTCACAGTACATATACCTTTCGGTGAAGGAGATGATGCGCGGACCAAGG 392
QY 374 TGATAAATGGGCAATCGGAAACCTCAAAAGCTCACTCAACAACTTCGTAATCTATG 433
DB 393 TGGCAGACTGGCTCCCGACGGCTCCCGTCCGTCGCTCCCAAGGCGTCAAGCCGAGC 452
QY 434 GAAAAATACCTCACTGCTGGGCAACAGCGCGGTGGAAAAAGCGGCTTGGCGTTGCGC 493
DB 453 TCTGGAAGCTCGCTTGGCCGCGCACAGCCGAGAGGCGCACAGCGCTTTCCTCGGCT 512
QY 494 TAGGCAATGGCGCAACATTAGACCATTCACGTTTTCAGCTTCAATGGAATTGATC 553
DB 513 TGGGCAACGCC-----AAGACCCAGCTAACCTTCTCCGCGCTCATGCGACTGACC 563
QY 554 CAGTCGACGAACTTAACAAATATATTAGAACCGATCCGATATTTAAAGTAAACCGG 613
DB 564 CCGTCGCGCGGCAACGGGAAGTCTCCAGCTCCAGCCCAATCTCTCACTACAGAGCGCT 623
QY 614 AATCTTTGAGCTGACATACCGGTTGCAATGTTGGGAAACCGAAGTGGAGCGAAGTGG 673
DB 624 CCTCTTCTGGCATGCGCATGCGGTGCTGTCATCTGCGACCGGCGTGGCGAAGAGAGA 683
QY 674 ACAAGTGATG---CCACCATGGCGGACCAAGGACTTAACCATGAGAGAGTTTACAAG 730

```

```

DB 684 AGAATATATTTCTTCCCTCCCTGCGACCAAGACGTGAACACGCCGAGTTCTACCGCG 743
QY 731 AGTGTAGGCGAGAAAGCCATTTGCTGCTCGGATTAACGACATATGATATGTTGG 790
DB 744 AGTGCAGGCGCGCTGCTACTTCTGTGACCAAGACATACGGGCACTGGAATGCTCG 803
QY 791 ACGATGATTTGCCCCGTTTGTGGGTTATGCGCGGTTGTATGTTAAGATGGGCAAA 850
DB 804 ACGACGACGCCGCCCAAGTTCAAT-----CACCTGCTCTGCAAGATGGGAAACG 851
QY 851 GAAAAAGTGTGATGAGGAGCTTTGTAGTGAATTTGTTGCTTCTCAAGTATA 910
DB 852 GTGCAAGGCGCAAGATCGGAGGTGCTGAGATCAATGATGAGGATTTCTTAATGCTG 911
QY 911 GTTGTGGGGTGAAGAAAGCGGATTCGATTTGATGAGATTCCTTCGTTTCTCGG 970
DB 912 CTTTGGGTGAGAAAGATGCAGATCTTGAAGCCATCTGAGAGACCGGCGGTTGCACCA 971
QY 971 CCAAGCTTGATCC 983
DB 972 CCAAGCTTGATCC 984
RESULT 11
AD38078
ID AD38078 standard; cDNA; 1104 BP.
XX
AC AD38078;
XX
DT 10-SEP-2002 (first entry)
XX
DE Soybean chlorophyllase cDNA from sfln1.pk002.m10:fls clone.
XX
KW Soybean; chlorophyllase; chlorophyll degradation; plant cell senescence;
XX enzyme; gene; ss.
XX
OS Glycine max.
XX
FH Key Location/Qualifiers
FT CDS 26..967
FT /tag= a
FT /product= "Soybean chlorophyllase protein from
FT sfln1.pk002.m10:fls clone"
FT /EC_number= "3.1.1.14"
XX
XX W0200229022-A2.
XX
PD 11-APR-2002.
XX
XX 04-OCT-2001; 2001WO-US031059.
XX
XX 05-OCT-2000; 2000US-0238161P.
XX
XX (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
XX Cahoon EB, Cahoon RE, Thorpe CJ;
XX
XX WPI; 2002-444102/47.
XX
XX P-PSDB; AAE23780.
XX
XX An isolated polynucleotide encoding a plant chlorophyllase polypeptide,
PT used to produce transformed plants that have controlled induction or
PT postonement of senescence.
PS Claim 5; Page 53; 69pp; English.
XX
XX The invention relates to plant chlorophyllase (EC 3.1.1.14) proteins and
CC polynucleotides encoding such proteins. Chlorophyllase enzyme is involved
CC in chlorophyll degradation in plant cell senescence. Sequences of the
CC invention are used to produce a plant by transforming a plant cell with
CC chlorophyllase and regenerating a plant from the transformed plant cell.
CC They may also be used to transform cells. The plants that are produced

```

CC can have economical importance as they can allow for the controlled
 CC induction or postponement of senescence. The present sequence is soybean
 CC chlorophyllase cDNA from s11n1.pk002.m10.f15 clone
 XX
 SQ Sequence 1104 BP; 308 A; 225 C; 249 G; 322 T; 0 U; 0 Other;

Query Match 16.4%; Score 194.4; DB 6; Length 1104;
 Best Local Similarity 56.3%; Pred. No. 6.6e-48;
 Matches 473; Conservative 0; Mismatches 346; Indels 21; Gaps 5;

```

QY 154 AATGATTCAACAGCAGCCGCAAAACCGGTGAGATACCTGTCCAAACAGTCGCGGAAT 213
DB 113 AGTGCCTCTCTCTCACTCCAAACCATTTGCTAATTTTACCACTGTTCTTGCGCA 172
QY 224 TATCCGCGTATTATTTCTCCATGCTTTATCTTCGCACTACTTCTACTCTGAGCTT 273
DB 173 TACCTGTATATTTGTGTCCATGCTTTTCATTTCCGCAATTTCTACTCAAAAGCTC 232
QY 274 CTTAACACATCGCTTCGCATGTTACATTTCTTGAGCCCC--ACAGTTGCAATTA 330
DB 223 CTAGCCCACTATGCTCAGATGATTCATATCGTCTCTCCCACTGTTTCCAAATGG 292
QY 331 TTGCGCGCGGAGGCGCAAGTGAAGTGAACATGCTGGAAGTGTATTAACCTGGCATCG 390
DB 293 CTTCTATGATGAGACCACTGAAGTGAATATGAGAAAAGTTGCGGATGATAGCT 352
QY 391 GAAAACCTCAAGCTCAGCTACCACTTCGTAATATGTAATGAAAATACACCTCATC 450
DB 353 GAGGAGCTTCAACATTTGCTTCCAGAAAGCTTGAAGCAATTTGGACAAACTGTTCTA 412
QY 451 GTGGCCACAGCCGCGTGGGAAAACGCGTTCGCTGCGCTAGGCCATGCCGCAACA 510
DB 413 TCAGGTCACAGTAGGGGTGGGAAAACGTGATTTGTGTGGCTTGTGCTATC----- 465
QY 511 TTAGACCATTCATCACTTTCACTCTAATAGAAATGATTCAGTCGAGAACTATAC 570
DB 466 -AAAACTAATCTCAAGTTTCAACACTTGTAGGCAATAGACCTGTGGCTGGCATCT 523
QY 571 AAATCATTTAGAACCGATCCGATATCTTAACGTATTAACCGGAATCTTTGAGCTGAC 630
DB 524 AAATATTTGAGAACGTCCTCATATATCTCACTGGCAAGCCACGGTCTTTGATTTGAAA 583
QY 631 ATACCGGTTGACAGTGTGGGAAACCGGACTCGGAC--GAAGTGAACACGTAATGCA 687
DB 584 ATGCGAGTTGAAGTATTTGGCACTGGATTTGGGCCAGAGAAAGCTTAATTTGTTACTCA 643
QY 688 CCATGCGCAACCAACGCACTTAAACCATGAGAGTTTACAAAGAGTGTAAAGCGACGAAA 747
DB 644 CCGTGTGCTCTGATGGGGTGAACATAAGGAGTTCTTCAACGAGTGCMAACCCCTTGT 703
QY 748 GCCCATTTGCGGCGGATTAAGGACATATGATATGATTTGAGAGATTTGCCGGT 807
DB 704 GCTAATTTGTTGACAAAGTATGATCATGACATGATGTAATGATGACACACAGAGG 763
QY 808 TTTTGTGG3--TTTATGGCCGTTGTATGTGTAGAAATGGGCAAA--GAAAAAGTCT 861
DB 764 CTAATTTGGGACATTTGTGTCAAAAGTATGTGTAAAGATGGGACGAGCTCTAAGGAC 823
QY 862 GAGATGAGAGCTTTGTATGAGTGAATTTGTGTTGCGTTCTCAAGTATATTTGTGGGT 921
DB 824 TTGATGAGAAAGGACCACTGAGGGGTGTGTGCTTCTTGAGGGGCAAAATTTGATAC 883
QY 922 GAAAAAGGAGATTCATGATTTGTAAGATTCCTTCGTTTCTCCGCGCAAGTTGAT 981
DB 884 CTATGGAAGGATTTGATGCTATTTTAAAGACCTTAATCTTGTCCACATGAAAGTAT 943

```

RESULT 12
 AAD38079
 ID AAD38079 standard; cDNA; 1125 BP.
 XX
 AC
 XX AAD38079;
 XX

```

DT 10-SEP-2002 (first entry)
XX
XX Soybean chlorophyllase cDNA from s12.pk130.f15:fls clone.
DE
XX
XX Soybean; chlorophyllase; chlorophyll degradation; plant cell senescence;
KW enzyme; gene; ss.
XX
XX Glycine max.
XX
XX Key Location/Qualifiers
FT CDS 42..992
FT /*tag= a
FT /product= "Soybean chlorophyllase protein from
FT s12.pk130.f15:fls clone"
FT /EC_number= "3.1.1.14"
FT
FT
FN MO200229022-A2.
FN
FN 11-APR-2002.
FN
FN 04-OCT-2001; 2001MO-US031059.
FN
FN 05-OCT-2000; 2000US-0238161P.
FN
FN (DUPO ) DU PONT DE NEMOURS & CO E I.
FN
FN Cahoon EB, Cahoon RE, Thorpe CJ;
FN
FN WPI; 2002-444102/47.
FN
FN P-PSDB; AAE23781.
FN
FN An isolated polynucleotide encoding a plant chlorophyllase polypeptide,
PT used to produce transformed plants that have controlled induction or
PT postponement of senescence.
PT
XX
XX Claim 5; Page 55; 69pp; English.
XX
XX The invention relates to plant chlorophyllase (EC 3.1.1.14) proteins and
CC polynucleotides encoding such proteins. Chlorophyllase enzyme is involved
CC in chlorophyll degradation in plant cell senescence. Sequences of the
CC invention are used to produce a plant by transforming a plant cell with
CC chlorophyllase and regenerating a plant from the transformed plant cell.
CC They may also be used to transform cells. The plants that are produced
CC can have economical importance as they can allow for the controlled
CC induction or postponement of senescence. The present sequence is soybean
CC chlorophyllase cDNA from s12.pk130.f15:fls clone
XX
XX
SQ Sequence 1125 BP; 315 A; 219 C; 262 G; 329 T; 0 U; 0 Other;
XX
Query Match 15.8%; Score 188; DB 6; Length 1125;
Best Local Similarity 56.6%; Pred. No. 5.8e-46;
Matches 479; Conservative 0; Mismatches 340; Indels 27; Gaps 6;

```

```

QY 154 AATGATTCAACAGCAGCCGCAAAACCGGTGAGATACCTGTCCAAACAGTCGCGGAAT 213
DB 132 ACTGCTTCTCTCACTCCAAACCATTTGCTAATTTTACCACTGTTCTTGCGCA 191
QY 214 TATCCGCGTATTATTTCTCCATGCTTTATCTTTGCAACTCTTCTACTCTGAGCTT 273
DB 192 TACCTGTATATTTGTTTTCGATGCTTTTGTATTTGCACTAGCTACTACTTAACCTC 251
QY 274 CTTAACACATCGCTTCGCATGTTACATTTCTTGAGCCCCACA--GTTGTGCAAAATTA 330
DB 252 CTAGCCCACTATGTTTACATGATTCATACTTGTCTCTCAAGCTGTTTCCATTTGGG 311
QY 331 TTGCGCGCGGAGGCGCAAGTGAAGTGAACATGCTGGAATGTGATTAACCTGGCATCG 390
DB 312 GTGCTATGTTTGGACCAAGAAAGTAAAGTGAAGAAAGTGTGATTTGGCTAGAT 371
QY 391 GAAAACCTCAAGCTCAGCTACCACTTGGTAAATGCTAATGAAATAATACACTCATC 450
DB 372 AACGGGCTTCAACCATTTGCTTCCGAGAGCGTTGAAGCCAAACTGAGAAACTGTTCTA 431

```

QY 451 GTGGGCGACAGCCGGGCGGAGAAAACGGCGTTTGGCGTTAGGCCATCCGCCACA 510
 DB 432 GTAGTCCAGAGAGGGTGGAAAACAGCAATTTGCTGGGCACTTGGTTA-----C 482
 QY 511 TTAGACCCATCATCAGCTTTTCACTCTTAATAGGAATTCAGTCGACAG-----A 564
 DB 483 TGTAAACCAAGCTCAAGTTTTCAGCACTCATAGGCATAGATCTCTGGTGGGTATCA 542
 QY 565 ACTAACAAATATCATTTAGAACCGATCCGATATCTTAACGTAATTAACCGGAATCTTTGAG 624
 DB 543 AAGGTAAAGCCTTGTCCATCTTCTGATATCTCAAGAGTGGCCAGGCTCTTAT 602
 QY 625 CTGACATATCCGTTGCGATGCTGGGAAACCGGACTCCGAC--GAATGGAACAACGTG 681
 DB 603 CTGAACATACCTGTGCTGTATATGGAACCTGGTGGGCCAGAGAAAGCTAATTTCTTT 662
 QY 682 ATGCGACCATGCGGACCAACGGAATTAACCATAGAGAGTTTCAAAAGATGTAAAGCG 741
 DB 663 TTTCCACCATGTGCTCCAAATGGGGTGAACATAAAGTTTCTCTGAGTGCACCA 722
 QY 742 ACGAAAGCCATTTCTGCTGCTGGCGATTAACGACATATGATATGTTGACGATGATTG 801
 DB 723 CCTATGCTATTTTGTGTGCAACGATATATGTCACATGATGATGATGATGATAA 782
 QY 802 CCCGTTTGTGGG---TTATAGCCGCTGTATATGTAAGATGGGCAAA---GAAA 855
 DB 783 CCAAGGCTAATTTGGACATATATGCAAGTATATGCAAGAAATGGAAAGAGGTCCT 842
 QY 856 AAGTCTGATGAGAGAGCTTTGTAGTGAATTTGTTGCTTCAATATAGTTTG 915
 DB 843 AGGACTTTGATGAGAGAGCTGTGGAGGTTGTTGCTTCTTGAGACACAGTTG 902
 QY 916 TGGGCTGAAAAAGCGAGATTCATGATTTGTAAGATCTTCCGTTCTCCGCCAAG 975
 DB 903 AATAGCGCTGGAAGATTTGATCTATTTAGCGAGTCTAATCTATCTCCGCCAAG 962
 QY 976 CTGAT 981
 DB 963 CTGAT 968
 RESULT 13
 AAD38080
 ID AAD38080 standard; cDNA, 1444 BP.
 XX
 AC AAD38080;
 XX
 DT 07-AUG-2003 (revised)
 DT 10-SEP-2002 (first entry)
 XX
 DE Tuliip chlorophyllase cDNA from etp1c.pk005.d16:file clone.
 XX
 KM Tuliip; chlorophyllase; chlorophyll degradation; plant cell senescence;
 KM enzyme; gene; ss.
 XX
 OS Tuliipa geeneriana.
 XX
 FH Key Location/Qualifiers
 FT CDS 97..1107
 FT /*tag= a
 FT /product= "Tuliip chlorophyllase protein from
 FT etp1c.pk005.d16:file clone"
 FT /EC_number= "3.1.1.14"
 FT
 XX
 PN MO200229022-A2.
 PD 11-Apr-2002.
 XX
 PF 04-OCT-2001; 2001MO-US031059.
 XX
 PR 05-OCT-2000; 2000US-0238161P.
 XX
 PA (DUPO) DU PONT DE NEMOURS & CO E I.

XX
 FI Cahoon EB, Cahoon RE, Thorpe CJ;
 ,XX
 DR WPI; 2002-444102/47.
 DR P-PSDB; AAE23782.
 XX
 PT An isolated polynucleotide encoding a plant chlorophyllase polypeptide,
 PT used to produce transformed plants that have controlled induction or
 PT postponement of senescence.
 PS Claim 5; Page 56-57; 69pp; English.
 XX
 CC The invention relates to plant chlorophyllase (EC 3.1.1.14) proteins and
 CC polynucleotides encoding such proteins. Chlorophyllase enzyme is involved
 CC in chlorophyll degradation in plant cell senescence. Sequences of the
 CC invention are used to produce a plant by transforming a plant cell with
 CC chlorophyllase and regenerating a plant from the transformed plant cell.
 CC They may also be used to transform cells. The plants that are produced
 CC can have economical importance as they can allow for the controlled
 CC induction or postponement of senescence. The present sequence is tulip
 CC chlorophyllase cDNA from etp1c.pk005.d16:file clone. (Updated on 07-AUG-
 CC 2003 to correct OS field.)
 CC
 SQ Sequence 1444 BP; 460 A; 295 C; 298 G; 391 T; 0 U; 0 Other;
 Query Match 15.1%; Score 178.8; DB 6; Length 1444;
 Best Local Similarity 54.1%; Pred. No. 4.1e-43;
 Matches 438; Conservative 0; Mismatches 357; Indels 15; Gaps 3;
 QY 156 TGAATTCACAGCACCGCGCAAAACCGGAGAAATCACCTGTCACAAAGTCGCGGAATTA 215
 DB 195 TGTCTCCGACACACTGGAGGACCACTGATGCACTTCATTGAGAAAGCGAGTA 254
 QY 216 TCCCGTGTCTTATATCTTCATGAGCTTTTATCTTCGCACTACTCTGACGTTCT 275
 DB 255 CCGAAGCTTCTACTCTCTTATGATTCATGATTCATGCTTCAACACTTCTGAGCTTAT 314
 QY 276 TAACCATATGCTTGCATGCTTATCATTTCTTGAGCCCAAGTTGCAAAATTTATGCC 335
 DB 315 CCAACATATGCAATCCCATGCTTATATGTTGTTGTAACCTGATATATCTTGTAGCTAC 374
 QY 336 GCGGAGAGGCAAGTGAAGTGAAGTGTGGAAGTGTAACTGGGCAATCGGAA 395
 DB 375 ATGTGATATGACGAATGCAATCAAGTCTGCAAAACACAGATTTGTAAGATGG 434
 QY 396 CTTCAAGCTCACCTACCACTTCGTAATATGCTAATGAAAAATACACTCTCGTGG 455
 DB 435 ACTGCAAGATGTTCTCCCAACAAAGTCAACACAGACTTAAGAACTGGAGTCGCGG 494
 QY 456 CCAACGCGCGGTGGAAAAAGCGGCTTGGCGTTGGCGTATGCGCATGCGCAACATTAGA 515
 DB 495 ACATAGCCCTGGGCGGCAAAAGATGATTTGCTCTTGGCACTGAGATATGC-----GAA 545
 QY 516 CCATCCATCAGCTTTACGCTCTAATAGGAATTTGATTCAGTGGCAAGAACTAAATA 575
 DB 546 GACTCAATTAAAGCTTCTCAAGCTGATCGGATGATCTGTGACGAGATACGAAAGGG 605
 QY 576 CATTAAACCGATCCGATATCTTAAAGTAAACCGGAATCTTTGAGCTGACATACC 635
 DB 606 ACAACAAACCAATCTCTGTACTGAAATATATCCCTCATCTTTGGAATCAAGATGCC 665
 QY 636 GGTTCAGTGTGGGAACCGGACTCGGACGAATGGAACACAGTATGCCACCATGCGC 695
 DB 666 GTCAATTAATATCGGAACAGGTTAAGTGAATGAAGAGAACTATTT--TGCCTGTGC 722
 QY 696 ACCAAGCACTTAACCATGAGAGTTTAAACAAGATGAAGCGGACGAAGCCCATTT 755
 DB 723 CCTTAAGGAGTAATATCAACCAAGATTTTATGATGATGTTCTTCCCGCTTGGCAATTT 782
 QY 756 CGTGGCTGCGATTAACGACATATGATATGTTGACAGATGATTTCCCGTTTGTGG 815
 DB 783 TGTGCTAAGACTATGATGTCATGACATGACATGATGATGATGATGATGACGACCTAGAGG 842

Qy	816	GTTTATGCGCCGTTGATGTGTATGAGANTGGCAAGAAAAAGTCAGATGAGACCTT	875
Db	843	TATTGTCTCACTACGATGCTGCAAAATGGAATCAAGAA-----CACCCATGCACGT	899
Qy	876	TGTATGATGGAATTGTGTGTTGCCGTTTCTCAAGTATGTTGTGGGGTGAAGAGCGAGAT	935
Db	900	TGTAGGGGAGATGCTGTGCTTCTTCTGAAGCCGCTTACAAGGTGATTAACAGATTT	959
Qy	936	TCGATTGATGTGAAGATCCCTCCGTTTC	965
Db	960	GCGAGCTCTGAAGGATTAACCCGAGATTTC	989

RESULT 14

XX	AD38075 standard; cDNA; 1216 BP.
XX	AAD38075;
DT	10-SEP-2002 (first entry)
XX	Grape chlorophyllase cDNA from vrilc.pk008.021.fis clone.
DE	Grape; chlorophyllase; chlorophyll degradation; plant cell senescence;
KM	enzyme; gene; ss.
XX	
OS	Vicia sp.
XX	
FH	Key Location/Qualifiers
FT	CDS 3..938
FT	/*tag= a
FT	/product= "Grape chlorophyllase protein from
FT	vrilc.pk008.021.fis clone"
FT	/EC_number= "3.1.1.14"
FT	/note= "CDS does not include start codon"
FT	/partial
XX	
PN	M0200229022-A2.
PD	11-APR-2002.
PF	04-OCT-2001; 2001MO-US031059.
XX	
PR	05-OCT-2000; 2000US-0238161P.
XX	
PA	(DUPO) DU PONT DE NEMOURS & CO E I.
P1	Cahoon EB, Cahoon RE, Thorpe CJ;
DR	WPI; 2002-444102/47.
XX	
XX	P-PADB; AAE23777.
PT	An isolated polynucleotide encoding a plant chlorophyllase polypeptide,
PT	used to produce transformed plants that have controlled induction or
PT	postponement of senescence.
XX	
PS	Claim 5; Page 48-49; 69pp; English.
XX	
CC	The invention relates to plant chlorophyllase (EC 3.1.1.14) proteins and
CC	polynucleotides encoding such proteins. Chlorophyllase enzyme is involved
CC	in chlorophyll degradation in plant cell senescence. Sequences of the
CC	invention are used to produce a plant by transforming a plant cell with
CC	chlorophyllase and regenerating a plant from the transformed plant cell.
CC	This may also be used to transform cells. The plants that are produced
CC	can have economical importance as they can allow for the controlled
CC	induction or postponement of senescence. The present sequence is grape
CC	chlorophyllase cDNA from vrilc.pk008.021.fis clone
XX	
Q0	Sequence 1216 BP; 331 A; 275 C; 256 G; 354 T; 0 U; 0 Other;

Query Match	14.3%;	Score 169.8;	DB 6;	Length 1216;
Best Local Similarity	53.2%;	Pred. No. 2.1e-40;		
Matches 435; Conservative	0;	Mismatches 367;	Indels 15;	Gaps 3;

QY	167	CACCGCCAAAACCGGAGAAATACCTGTCACAAAGTGGCGGAACTTATCCGTCGTT	228
Db	112	CCCCCACACGAACTCTGAATGTGACGCCCTCGAAGCAGGGAGTTCCTCCGTCGTC	171
QY	227	TATCTTCATGGCTTTTATCTTCGCAACTACTTCTACTGACGGTCTTAAACACATCG	288
Db	172	TCCTCCTCATGGTTATCTTCTCTAATACTCTTTCTACTCCACGCTATCCAACTAG	233
QY	287	CTTCGATGGTTACATTTCTTGAGCCCAAGTGTGCAAAATATGTGCCGCGGAGGC	348
Db	232	CCCTCATGGTTCAATGTCTTGCTCCCTCAGTTTATACATGTGGCTGACCAATTCMA	291
QY	347	AAGTGAAAGTGCATGTGTGAAGTGTAACTGGGCAATCGGAAACCTCAAACTC	408
Db	292	GCGAAGAGTCAAGTCCGACGCTCTTAACTAAATGGTTATCCAAAGACTCCATAC	351
QY	407	ACCTACCAACTTGGGTAAATGCTAATGCAAAATACACTCACTCGTGGGACAGCCGC	468
Db	352	TACTTCCTCCCATGTGTGGCCAAATTTAAGCAACTAGACTTGGCGGCATGTGTG	411
QY	467	GTGGGAAACAGCGCTTGGCGTTGGCGCTAGCCATGCGCAACTTGAACCATTCATCA	528
Db	412	GAGGCAAAATGCTTTTGCTTAGACACTGAG-----AAAGATCCCACTTCTGTGA	462
QY	527	CGTTTCAGCTCAATAGGAATTGATCCAGTCGACGAACTTAACTAATCATTTAGAACG	588
Db	463	AATTTTCAGCTTGATAGGCATAGACCCGGTGTGATGTAAATGACAAAGGAAACAAACC	522
QY	587	ATCCGCATATCTTAAAGTATTAACCGGAATCTTTGACTGGAACATACCGTTGCAATG	648
Db	523	CTCACACGGTACTCACCTATGTTCCATTTGATCTTAGACATGGCAATGATGTAA	582
QY	647	TGGGAACCGGACTCG--ACGGAATGTGAACAAGTATGCCACATGGCCACCAACGG	703
Db	583	TTGGTTCGGGTTGGGTGAAGTGAAGGAACCCCTGTCTCCCTCTTGTCGCCCAAGG	642
QY	704	ACTTAAACCATGAGAGTTTACAAAGAGTGTAAAGGACAGCAAAAGCCATTTGCTGCTG	763
Db	643	GCGTAAACCATGAGACTTCTTAAAGATGCCGGAACAGCTGTTATTTTCTTGCCA	702
QY	764	CGGATTAACGACATATGATATGTGACAGATGATTTGCCGTTTTGTGGTTTATGG	823
Db	703	AGGACTATGGCCATCTTGACATGTAGACAGATGACATAATGAAATTTAGAGGAAACCTA	762
QY	824	CCGGTTGTATGTGTAAAGATGGGCAAAAGAAAAGTGTGAATAGAGACTTTTAGTG	883
Db	763	CACATGTGTGTGTAAAAATGGAAGTCTAGAAACC---CATGAGAGGTTTGTGTGAG	813
QY	884	GAATGTGGTTGGCTTCTCAAGTATAGTTTGTGGGTGTAAGAAAAGCGAGATGTGATTTGA	943
Db	820	GCATTTGATGTGATTTATGAAGAGTATTTTGAAGGCGGATACAGCAGCTTATCTCCA	879
QY	944	TTGTGAAGATCTTCCGTTTCTCCGGCCCAAGCTTGA	980
Db	880	TTAGAGATGGGCATGCTACTGCACAGTGGAGCTTCA	916

RESULT 15

Accession	Gene	Accession	Gene
AAH41137	Arabidopsis thaliana chlorophyllase coding sequence #2	AAH41137	standard; DNA, 1135 BP.
ID			
XX			
XX			
AC	AAH41137;		
XX			
XX			
DT	21-AUG-2001 (first entry)		
XX			
XX			
DE	Arabidopsis thaliana chlorophyllase coding sequence #2		
XX			
XX			
KW	Chlorophyllase; transgenic plant; ds.		
XX			
OS	Arabidopsis thaliana.		
XX			
PN	JP2001086990-A.		

```
XX 03-APR-2001.
BD 20-SEP-1999; 99JP-00266181.
PF 20-SEP-1999; 99JP-00266181.
PR 20-SEP-1999; 99JP-00266181.
XX (KAGO ) KAGOME KK.
XX WPI; 2001-338421/36.
DR P-PSDB; NAB99103.
XX
XX DNA encoding chlorophyllase, useful for producing transgenic plants.
XX
XX Claim 5; Page 12-14; 21pp; Japanese.
XX
XX The present sequence is a chlorophyllase protein coding sequence. The
XX chlorophyllase protein coding sequence can be used for the transformation
XX of a plant
XX
SQ Sequence 1135 BP; 331 A; 233 C; 247 G; 324 T; 0 U; 0 Other;

Query Match      13.8%; Score 163.4; DB 5; Length 1135;
Best Local Similarity 54.3%; Pred. No. 1.8e-38;
Matches 420; Conservative 0; Mismatches 341; Indels 12; Gaps 4;

QY 157 GATTCAACAGCAGCCGAAACCGGTGAAATCACTGTCCAGAGTCGCGAATTAT 216
DB 151 GCTTCACCGTCTCCGCAAGACAGCTGTGGCTACGCGGTGAGAGAGATTAT 210
QY 217 CCGGTGTTTATCTTCAGTGGCTTTATCTGCAACTACTCTGACGTTCT 276
DB 211 CCGGTGTTTATCTTCAGTGGCTTTATCTGCAACTACTCTGACGTTATG 270
QY 277 AACCAATGCTGCTGCAATGTTAATCTGTAGCCCAAGTTGCAATTATGCG 336
DB 271 TTGATGTTCTTCTCATGCTTCTCATGCTCTCATGTTATATAGTATGCGGA 330
QY 337 CCGGAGGGGCAATGGAAGTGAAGCTGGAAGTGTATAACTGGGACTCGAANA 396
DB 331 CCAGACACATGATGATTAATCAACGCGGAGATTATGATTGTTATCAGTAGA 390
QY 397 CTCAAGCTCACTACCAACTTCGTTAATGTAATGGAATACTCACTGCTGAGC 456
DB 391 CTAAATCACTTCTTCAGCGCAAGTAAACCAAACTATCAAAATTTGCCCTCCG 450
QY 457 CACAGCGCGGTGGGAAACCGCGTTGCGGTGCGCTAGGCAATGCCAATAGAC 516
DB 451 CATAGCGCGGTGGGAAACCGCGTTGCGGTGCGCTTAAGAAATTG---GGTACTCC 507
QY 517 CCATTCATCACTTTTCACTCTAATAGAAATTGATCCAGTGCAGAACTAACAAATAC 576
DB 508 TCGAATCTAAAGATCTCGACATGATCGTATGATCCAGTGAACAGGAAAGGG 567
QY 577 ATTGAACCGATCCGCAATTTTAACGTAATAACCGAATTTTCGAGCTGGAATACCG 636
DB 568 AAACAAACCCCTCTCCGCTGTGCTTAACCTTCAAACTATTGACTAGACAAACG 627
QY 637 GTTGCA---GTGTGGGAACCGGACTCGG---ACCGAAGTGAACAACGTATGCCACA 690
DB 628 CCTATATCTGTGATCGGTGGGGCTTGTGAACCGCTCGGAACCATTAATCCACCG 687
QY 691 TGGCAGCCAGCGACTTAACATGAGAGTTTACAAAGAGTAAAGCGACGAAAGCC 750
DB 688 TGTACACCTCCCGAGTGAATCACCGAAGTTCCTTCGGAATGTCAAGGTCCAGCATGG 747
QY 751 CATTTCGTGCTGGGATTACGACATATGATATGTTGACGATGATTTGCCGGTTT 810
DB 748 CATTTCGTGCGAAGATTATGGGATTGTGACATGCTTGTATGATACAAAGGATTT 807
QY 811 GTTGGATTATAGCGCGTTGTATGTATGAATGGGCAAGAAAAAGTCTGAGATGAG 870
DB 808 AGAGGAAAGAGTTCTTATTTGTTGTATGAATGTGA---AGAGAGAGACCAATGAGG 864
```

```
QY 871 AGCTTTAGGTGGAATTTGTGCTGCTTCTCAAGTATAGTTTGGGGTGA 923
DB 865 AGATTCGTGTGGACTGTGTATCATTTTGTGAAGGCTTATTTGGAAGGAGA 917
```

Search completed: March 20, 2006, 14:14:20
Job time : 704.094 secs

THIS PAGE LEFT BLANK

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocelebration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 20, 2006, 14:06:13 ; Search time 1101.06 Seconds

(without alignments)
8922.325 Million cell updates/sec

Title: US-10-634-548-18

Perfect score: 1188
Sequence: 1 gatcacataaattctcaacac.....cccccaaaaaaaaaaaaaa 1188

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 413468905 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA Main:*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
6: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	1188	100.0	1188	7	US-10-634-548-18
2	975	82.1	975	3	US-09-938-842A-158
3	975	82.1	975	3	US-09-938-842A-158
4	748	63.0	1125	7	US-10-425-114-26342
5	239.4	20.2	987	9	US-10-381-123-1
6	202.6	17.1	1174	9	US-10-381-123-7
7	199	16.8	1242	9	US-10-381-123-17
8	194.4	16.4	1104	9	US-10-381-123-9
9	188	15.6	1307	7	US-10-381-123-11
10	185.2	15.6	1307	7	US-10-424-599-45715
11	178.8	15.1	1444	9	US-10-381-123-13
12	169.8	14.3	1216	9	US-10-381-123-3
13	163.4	13.8	1135	7	US-10-634-548-19
14	135.6	11.4	1068	7	US-10-424-599-34940
15	110.8	9.3	1156	9	US-10-381-123-25
16	109.8	9.2	635	7	US-10-424-599-131410
17	109.6	9.2	796	7	US-10-424-599-86252
18	109.2	9.2	751	8	US-10-425-115-70835
19	109.2	9.2	1302	9	US-10-381-123-5
20	109.2	9.2	2439	8	US-10-424-599-140359
21	107.8	9.1	538	7	US-10-424-599-34941
22	106.6	9.0	1244	7	US-10-425-114-25386
23	106.6	9.0	1244	7	US-10-425-114-25387

24	98.6	8.3	916	9	US-10-381-123-15	Sequence 15, Appl
25	82.6	7.0	1236	8	US-10-425-115-166593	Sequence 166593,
26	82.6	7.0	1274	9	US-10-381-123-27	Sequence 27, Appl
27	77.8	6.5	1004	7	US-10-424-599-68342	Sequence 68342, A
28	56.6	4.8	408	7	US-10-437-963-80160	Sequence 80160, A
29	53	4.5	1155	7	US-10-437-963-94232	Sequence 94232, A
30	52.6	4.4	728	7	US-10-767-701-3821	Sequence 3821, Ap
31	47	4.0	965	3	US-09-938-842A-2858	Sequence 2858, Ap
32	47	4.0	985	3	US-09-938-842A-2858	Sequence 2858, Ap
33	45	3.8	1829	3	US-09-974-879-60	Sequence 60, Appl
34	45	3.8	1829	3	US-09-305-736-60	Sequence 60, Appl
35	45	3.8	1829	3	US-09-818-683-60	Sequence 60, Appl
36	45	3.8	1829	3	US-09-818-683-60	Sequence 60, Appl
37	45	3.8	1829	7	US-10-621-401-60	Sequence 60, Appl
38	41.6	3.5	631	3	US-09-814-353-5993	Sequence 5993, Ap
39	41.6	3.5	631	3	US-09-814-353-12272	Sequence 12272, A
40	40.8	3.4	13376	6	US-10-311-455-556	Sequence 556, App
41	40.4	3.4	524	7	US-10-767-701-3830	Sequence 3830, Ap
42	39.8	3.4	752	3	US-09-814-353-18656	Sequence 18656, A
43	39.6	3.3	542	4	US-09-925-065A-175209	Sequence 175209,
44	39.6	3.3	604	5	US-10-106-698-1620	Sequence 1620, Ap
45	39.6	3.3	1345	8	US-10-723-860-5698	Sequence 5698, Ap

ALIGNMENTS

RESULT 1
US-10-634-548-18
Sequence 18, Application US/10634548
Publication NO. US20040045051A1
GENERAL INFORMATION:
APPLICANT: No. US20040045051A1r1s, Susan R
APPLICANT: Lincoln, Kim
APPLICANT: Abad, Mark Scott
APPLICANT: Eilers, Robert
APPLICANT: Hartenberger, Karen Kindle
APPLICANT: Hirsberg, Joseph
APPLICANT: Karunanandaa, Balasubramini
APPLICANT: Moshiri, Farhad
APPLICANT: Stein, Joshua C.
APPLICANT: Valentin, Henry E.
APPLICANT: Venkatesh, Tyamagondlu V.
TITLE OF INVENTION: Tocopherol biosynthesis related genes and used thereof
FILE REFERENCE: Ren-01-125
CURRENT APPLICATION NUMBER: US/10/634,548
CURRENT FILING DATE: 2003-08-05
PRIOR APPLICATION NUMBER: us 60/400,689
PRIOR FILING DATE: 2002-08-05
NUMBER OF SEQ ID NOS: 79
SOFTWARE: PatentIn version 3.2
SEQ ID NO 18
LENGTH: 1188
TYPE: DNA
ORGANISM: Synechocystis PCC6803
US-10-634-548-18

Query Match 100.0%; Score 1188; DB 7; Length 1188;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	GATGACATAATCTTCAACACACTCTTTAATTACTGATTATATCAATATGCGGGGATTA	60
DB	1	GATGACATAATCTTCAACACACTCTTTAATTACTGATTATATCAATATGCGGGGATTA	60
QY	61	GAGGACAGTCACAGTTTCTGTCGTGTAATCCGCGCGCTTTGAGATGACAGCTTC	120
DB	61	GAGGACAGTCACAGTTTCTGTCGTGTAATCCGCGCGCTTTGAGATGACAGCTTC	120
QY	121	CCGACACCGAGATATACCGGTGATCCGGTGAATAATGATTCACACACCGCCAAACCG	180
DB	121	CCGACACCGAGATATACCGGTGATCCGGTGAATAATGATTCACACACCGCCAAACCG	180

181 GTGAGATACCTGTCACAAAGTCGCGGAACTTATCCCGTCTTTATCTTCATGGC 240
181 GTGAGATACCTGTCACAAAGTCGCGGAACTTATCCCGTCTTTATCTTCATGGC 240
241 TTTTATCTTCGCACTACTTCTTCTGACGTTCTTAAACCATGCTTTCGATGGTAC 300
241 TTTTATCTTCGCACTACTTCTTCTGACGTTCTTAAACCATGCTTTCGATGGTAC 300
301 ATCTGTGAGCCGCAAGTGTGCAATTAATTTCCGCGGAGGGCAAGTGAAGTGAAC 360
301 ATCTGTGAGCCGCAAGTGTGCAATTAATTTCCGCGGAGGGCAAGTGAAGTGAAC 360
301 ATCTGTGAGCCGCAAGTGTGCAATTAATTTCCGCGGAGGGCAAGTGAAGTGAAC 360
361 GATGCTGGAAGTGTGATTAACCTGCGCATCGGAAACCTCAAGCTCACCTCAACTTCG 420
361 GATGCTGGAAGTGTGATTAACCTGCGCATCGGAAACCTCAAGCTCACCTCAACTTCG 420
421 GTAAATGCTAATGAAAAATACCTCTCACTCGTGGGCAACAGCGCGGTGGAAAAAGCGCG 480
421 GTAAATGCTAATGAAAAATACCTCTCACTCGTGGGCAACAGCGCGGTGGAAAAAGCGCG 480
481 TTTGGGTTGGGCTGAGGCAATGCGCAACATTAGACCATCATGACGTTTTCAGCTCTA 540
481 TTTGGGTTGGGCTGAGGCAATGCGCAACATTAGACCATCATGACGTTTTCAGCTCTA 540
541 ATAGGAATTTGATCCAGTCGACAGGAATTAACAATATCATTTAGAACCGATCCGATATCTTA 600
541 ATAGGAATTTGATCCAGTCGACAGGAATTAACAATATCATTTAGAACCGATCCGATATCTTA 600
601 ACGTATTAACCGGAATCTTTTGAAGTGAATACCGGTTGCAAGTGTGGAAACCGGATCTC 660
601 ACGTATTAACCGGAATCTTTTGAAGTGAATACCGGTTGCAAGTGTGGAAACCGGATCTC 660
661 GGACGGAAGTGAACAAAGTGAATGCGCAATGCGCAACGGAATTAACAATGAGGAG 720
661 GGACGGAAGTGAACAAAGTGAATGCGCAATGCGCAACGGAATTAACAATGAGGAG 720
721 TTTTCAAAAGTGTGAGGCGCAAGAAACCATTTCTGAGTGTGCGGATTAACGATATG 780
721 TTTTCAAAAGTGTGAGGCGCAAGAAACCATTTCTGAGTGTGCGGATTAACGATATG 780
781 GATATGTTGAGAGATGATTTTGGCGGTTTGTGGGTTTATGCGCGGTTGATGTATG 840
781 GATATGTTGAGAGATGATTTTGGCGGTTTGTGGGTTTATGCGCGGTTGATGTATG 840
841 AATGGGCAAGAAAAAGTGTGAGATGAGGAGCTTGTAGGATGAAATGTGGTTGCGTT 900
841 AATGGGCAAGAAAAAGTGTGAGATGAGGAGCTTGTAGGATGAAATGTGGTTGCGTT 900
901 CTCAAGTATGATTTGTGGGTTGAAAAAGCGAGATTCGATTTGATGAGGATCTTCC 960
901 CTCAAGTATGATTTGTGGGTTGAAAAAGCGAGATTCGATTTGATGAGGATCTTCC 960
961 GTTTCTCCGCGCAAGCTTGTATCTTCACTGAGTGTGAAGAGCTTGTGATCTTCTCC 1020
961 GTTTCTCCGCGCAAGCTTGTATCTTCACTGAGTGTGAAGAGCTTGTGATCTTCTCC 1020
1021 TAGATTTGTGTATGATATGATATGAGAGGAGCTTGTAAATTTGAAAAACCTATCAATG 1080
1021 TAGATTTGTGTATGATATGATATGAGAGGAGCTTGTAAATTTGAAAAACCTATCAATG 1080
1081 TTTTCTAGCTCCAGCTAGCTATGTTTCAATGCTTAAGTGTGATATTTTATTTAA 1140
1081 TTTTCTAGCTCCAGCTAGCTATGTTTCAATGCTTAAGTGTGATATTTTATTTAA 1140
1141 CTGATCAAAACATTTGTATGTTTATGTTTACCCCAAAAAA 1188
1141 CTGATCAAAACATTTGTATGTTTATGTTTACCCCAAAAAA 1188

RESULT 2
US-09-938-842A-158
; Sequence 158, Application US/09938842A
; Patent No. US20020160378A1

GENERAL INFORMATION:
APPLICANT: Harper, Jeff
APPLICANT: Krepis, Joel
APPLICANT: Wang, Xun
TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
FILE REFERENCE: SCRIPI300-3
CURRENT APPLICATION NUMBER: US/09/938, 842A
CURRENT FILING DATE: 2001-08-24
PRIOR APPLICATION NUMBER: US 60/227, 866
PRIOR FILING DATE: 2000-08-24
PRIOR APPLICATION NUMBER: US 60/264, 647
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: US 60/300, 111
PRIOR FILING DATE: 2001-06-22
NUMBER OF SEQ ID NOS: 5379
SEQ ID NO 158
LENGTH: 975
TYPE: DNA
ORGANISM: Arabidopsis thaliana
US-09-938-842A-158

Query Match 82.1%; Score 975; DB 3; Length 975;
Best Local Similarity 100.0%; Pred. No. 3.2e-285;
Matches 975; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

49 ATGGCGGCGATAGAGACAGTCCAAAGTTTCTCTGTGTATCTCCGGGCGTTTGTAG 108
1 ATGGCGGCGATAGAGACAGTCCAAAGTTTCTCTGTGTATCTCCGGGCGTTTGTAG 60
109 ATAGGAGCTCTCCGCAACCGAGATACCGGTGATCCGGTGAATAATGATTCACAGCA 168
61 ATAGGAGCTCTCCGCAACCGAGATACCGGTGATCCGGTGAATAATGATTCACAGCA 120
169 CCGCAAAACCGGTGAGATACCTGTCAACAGTGCAGGAACTTATCCGTTGTTTA 228
121 CCGCAAAACCGGTGAGATACCTGTCAACAGTGCAGGAACTTATCCGTTGTTTA 180
229 TTTTCCATGAGCTTTTATCTTCCGATCTTCTGATCTGATCTGATCTTCAACATGCT 288
181 TTTTCCATGAGCTTTTATCTTCCGATCTTCTGATCTGATCTTCTGATCTTCAACATGCT 240
289 TCGATGTTTAACTTTCTGTAGCCCACTGTCACAAATTTATGCGCGGAGGAGCA 348
241 TCGATGTTTAACTTTCTGTAGCCCACTGTCACAAATTTATGCGCGGAGGAGCA 300
349 GTGAGATGAGCAGTGTGAGATGATTAACCTGAGCAATCGAAAACTTCAAGTCTAC 408
301 GTGAGATGAGCAGTGTGAGATGATTAACCTGAGCAATCGAAAACTTCAAGTCTAC 360
409 CTACCACTTCCGTAATGCTTAATGAAAAATACCTCACTGCGGCGCAACGCGGCT 468
361 CTACCACTTCCGTAATGCTTAATGAAAAATACCTCACTGCGGCGCAACGCGGCT 420
469 GGGAAAAAGCGTTTGGGTTGGCTGAGGAGTCCGCAACATTTAGACCATCCATCAG 528
421 GGGAAAAAGCGTTTGGGTTGGCTGAGGAGTCCGCAACATTTAGACCATCCATCAG 480
529 TTTTCACTCTAATGAGATTTGATCCAGTGCAGGAACTTAACAATATCATTTAGACCAT 588
481 TTTTCACTCTAATGAGATTTGATCCAGTGCAGGAACTTAACAATATCATTTAGACCAT 540
589 CCGGATATCTTAATGAGATTTGATCCAGGAACTTTTGAAGTGCATACCGGTTGCAAGTGTG 648
541 CCGGATATCTTAATGAGATTTGATCCAGGAACTTTTGAAGTGCATACCGGTTGCAAGTGTG 600
649 GGAACCGGATCTCGGACCGAGTGAACAGTGAATGCGCAATGCGGACCAACGAGCTTA 708
601 GGAACCGGATCTCGGACCGAGTGAACAGTGAATGCGCAATGCGGACCAACGAGCTTA 660
709 AACCATGAGAGTTTAAACAAGTGTAGGCGAGGAAACCCATTTCTGTGCTGCGGAT 768

Db	661	AACCATGAGAGTTTTCACAAAGATGTAAAGCGACGAAAGCCCATTTGTGCGCTCGGAT	720
Qy	769	TACGGAACATATGGAATATGTTGGACATATGATTTTCCCGGTTTGTGGGTTATAGCCCGGT	828
Db	721	TACGGACATATGGAATATGTGTGACCATATATTTGCCGTTTGTGTGGGTTATATGCCCGGT	780
Qy	829	TGTATGTGTAAAGATGGCCAAAGAAAAAGTCTGAGATGAGAGCTTTGTATGATGAAAT	888
Db	781	TGTATGTGTAAAGATGGCCAAAGAAAAAGTCTGAGATGAGAGCTTTGTATGATGAAAT	840
Qy	889	GTGTTGGGTTTTCACATATATGTTTGTGGGGTGA AAAAGCGGAATTCGATTAATGTG	948
Db	841	GTGTTGGGTTTTCACATATATGTTTGTGGGGTGA AAAAGCGGAATTCGATTAATGTG	900
Qy	949	AAGATATCTTCGTTTCTCGGCCCAAGCTGATTCCTTCACCTGAGTTGGAAGAACTTCT	1008
Db	901	AAGATATCTTCGTTTCTCGGCCCAAGCTGATTCCTTCACCTGAGTTGGAAGAACTTCT	960
Qy	1009	GGTATCTTCGTTCTAG	1023
Db	961	GGTATCTTCGTTCTAG	975

RESULT 3
US-09-93

Sequence 158, Application US/09938842A
Publication No. US20040009476A9

```

/ GENERAL INFORMATION:
/ APPLICANT: Harper, Jeff
/ APPLICANT: Kreps, Joel
/ APPLICANT: Wang, Xun
/ APPLICANT: Zhu, Tong
/ TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
/ TITLE OF INVENTION: SAME, AND METHODS OF USE
/ FILE REFERENCE: SCRIPI300-3
/ CURRENT APPLICATION NUMBER: US/09/938, 842A
/ CURRENT FILING DATE: 2001-08-24
/ PRIOR APPLICATION NUMBER: US 60/227, 866
/ PRIOR FILING DATE: 2000-08-24
/ PRIOR APPLICATION NUMBER: US 60/264, 647
/ PRIOR FILING DATE: 2001-01-16
/ PRIOR APPLICATION NUMBER: US 60/300, 111
/ PRIOR FILING DATE: 2001-06-22
/ NUMBER OF SEQ ID NOS: 5379
/ SEQ ID NO 158
/ LENGTH: 975
/ TYPE: DNA
/ ORGANISM: Arabidopsis thaliana
/ US-09-938-842A-158

```

Query Match	82.1%	Score 975	DB 3	length 975
Best Local Similarity	100.0%	Pred. NC	3.2e-265	
Matches 975	Conservative 0	Mismatches 0	Indels 0	Gaps 0

QY	49	ATGCGGCGAATAGAGACAGTCCAAAGCTTTCTCTGTGAGTAACTCCGGCGGCTTTGAG	108
Db	1	ATGGCGCGAATAGAGACAGTCCAAAGCTTTCTCTGTGAGTAACTCCGGCGGCTTTGAG	60
QY	109	ATAGGCAAGCCTCCGCAACAACCGAGTACCGGTGATCCGGTGAAATGATTCAACAGCA	168
Db	61	ATAGGCAAGCCTCCGCAACAACCGAGTACCGGTGATCCGGTGAAATGATTCAACAGCA	120
QY	169	CCGCGAAACCGGTGAGAAATCACTGTGTCCAAACAGTCCGCGGAACTTATCCGCTGTTTTA	228
Db	121	CCGCGAAACCGGTGAGAAATCACTGTGTCCAAACAGTCCGCGGAACTTATCCGCTGTTTTA	180
QY	229	TTCTTCACATGACTTTTATCTTCGCAACTACTTACTGTGACGTTCTTAAACAATCGCT	288
Db	181	TTCTTCACATGACTTTTATCTTCGCAACTACTTACTGTGACGTTCTTAAACAATCGCT	240
QY	289	TCCGATGGTTACATTTCTGTAGCCCAACAGTTGTGCAAAATTATTCGCCCGGAGGGGCA	348
Db	241	TCCGATGGTTACATTTCTGTAGCCCAACAGTTGTGCAAAATTATTCGCCCGGAGGGGCA	300

QY	349	GTGGAACTGGACGATGCTGGAAAGTGATATAA	CTGGGCAATCGGAAACCTCAAGCTCAC	408
Db	301	GTGGAACTGGACGATGCTGGAAAGTGATATAA	CTGGGCAATCGGAAACCTCAAGCTCAC	360
QY	409	CTACCAACTTCGGTAAATGCTAATGAAAAATAC	CACTCACTCTGTGGCCACAGCCGCGGT	468
Db	361	CTACCAACTTCGGTAAATGCTAATGAAAAATAC	CACTCACTCTGTGGCCACAGCCGCGGT	420
QY	469	GGGAAAAACGGGCTTTGGGGTTGGCGCTGAGG	CAATCGGCACATATGAGCCCATCCATCAGC	528
Db	421	GGGAAAAACGGGCTTTGGCGTTGGCGCTGAGG	CAATCGGCACATATGAGCCCATCCATCAGC	480
QY	529	TTTTCACTCTTAATATGGAATTTGATCCAATCG	CGAGAACTAAACAAATACATTAGAACCGAT	588
Db	481	TTTTCACTCTTAATATGGAATTTGATCCAATCG	CGAGAACTAAACAAATACATTAGAACCGAT	540
QY	589	CCGGAATCTTAAACGTATTAACCGGAATCTTT	CGAGCTGGAACATACCGGTTGCAGTGTGTG	648
Db	541	CCGGAATCTTAAACGTATTAACCGGAATCTTT	CGAGCTGGAACATACCGGTTGCAGTGTGTG	600
QY	649	GGAACCCGAGCTCGGACCCGGAAGTGGAA	CAACGTGATGCAACATGCGACCAACCGAATT	708
Db	601	GGAACCCGAGCTCGGACCCGGAAGTGGAA	CAACGTGATGCAACATGCGACCAACCGAATT	660
QY	709	AACCATAGAGACTTTTACAAAGAGTGTAAAG	CGACGAAAGCCATTTCGTGTGCTGCGGAT	768
Db	661	AACCATAGAGAGACTTTTACAAAGAGTGTAAAG	CGACGAAAGCCATTTCGTGTGCTGCGGAT	720
QY	769	TACGACATATGAGATATGTTGACGATGATTTG	CCCGGTTTGTGTGGGTTTATATGGCCGAT	828
Db	721	TACGACATATGAGATATGTTGACGATGATTTG	CCCGGTTTGTGTGGGTTTATATGGCCGAT	780
QY	829	TGTATGTGTAAAGATGGGCAAGAAAAAGTCTG	AGATGAGAGACTTTGTATGAGTGTGAATT	888
Db	781	TGTATGTGTAAAGATGGGCAAGAAAAAGTCTG	AGATGAGAGACTTTGTATGAGTGTGAATT	840
QY	889	GTGGTTGCGTTTCTCAAGTATAGTTTGTGGGG	TGAAGAAAAACGGAGATTTCATGATGTTGTG	948
Db	841	GTGGTTGCGTTTCTCAAGTATAGTTTGTGGGG	TGAAGAAAAACGGAGATTTCATGATGTTGTG	900
QY	949	AAGGATCTTTCGGTTTCTCCGGCCAACTTGAT	CTTTCACCTGAGTTGGAAGAAAGCTTCT	1008
Db	901	AAGGATCTTTCGGTTTCTCCGGCCAACTTGAT	CTTTCACCTGAGTTGGAAGAAAGCTTCT	960
QY	1009	GGTATCTTTCGCTTAG	1023	
Db	961	GGTATCTTTCGCTTAG	975	

RESULT 4
US-10-42

! Sequence 26342, Application US/10425114
! Publication No. US20040034888A1

```

1  GENERAL INFORMATION:
2  APPLICANT: Liu, Jindong
3  APPLICANT: Zhou, Yihua
4  APPLICANT: Kovalic, David K.
5  APPLICANT: Screen, Steven E
6  APPLICANT: Tabaska, Jack E
7  APPLICANT: Cao, Yongwei
8  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
9  TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
10 FILE REFERENCE: 38-21(5313)B
11 CURRENT APPLICATION NUMBER: US/10/425,114
12 CURRENT FILING DATE: 2003-04-28
13 NUMBER OF SEQ ID NOS: 73128
14 SEQ ID NO 26342
15 LENGTH: 1125
16 TYPE: DNA
17 ORGANISM: Brassica napus
18 FEATURE:
19 OTHER INFORMATION: Clone ID: LIB4156-010-D3_FLI

```



```

Db      617 CTGATATGCTCACTTGGACATGTTAGATGATCATCTTTCAGGCTGACTTGGGCGCATTT 676
Qy      824 CCGGTTGATGTATGATAGATGGGCAAAAAGCTGAGATGAGAGCTTTGTAGGTG 883
Db      677 GCGGTTATATCTGCAAGATGGGAAAGGCTCCTAGGAGCCCATGAGAGATGTGTGGTG 736
Qy      884 GAATTTGTGTCGTTTCTCAAGTATGTTTGTGGGGTGAAGAAAGCGGAGATTGATTTGA 943
Db      737 GCGTTTTGTGTGCAATCTTGAAGGCTTATTTGGAAGGTCAGACTGAGATTTCAAGCCA 796
Qy      944 TTGGAAGGATCTTCGCTTCTCGGCGCAAGCTTGATCT 984
Db      797 TTGTTGATGAACCTGATCTGGCTCCTGTGAAGCTTGATCT 837

```

RESULT 6

```

US-10-381-123-7
; Sequence 7, Application US/10381123
; Publication No. US20050081263A1
; GENERAL INFORMATION:
; APPLICANT: E.I. du Pont de Nemours and Company
; TITLE OF INVENTION: Chlorophyllases
; FILE REFERENCE: B01477 PCT
; CURRENT APPLICATION NUMBER: US/10/381,123
; CURRENT FILING DATE: 2003-03-21
; PRIOR APPLICATION NUMBER: PCT/US01/31059
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: 60/238,161
; PRIOR FILING DATE: 2000-10-05
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 7
; LENGTH: 1174
; TYPE: DNA
; ORGANISM: Glycine max
US-10-381-123-7

```

```

Query Match      17.1%; Score 202.6; DB 9; Length 1174;
Best Local Similarity 56.9%; Pred. No. 3.5e-50;
Matches 480; Conservative 0; Mismatches 339; Indels 24; Gaps 5;

```

```

Qy      154 AATGATTCACAGACGCGCAAAACCGGTGAGATCACTCTGCCAACAGTCCGCGAACT 213
Db      102 AATGCTTCCTCTCACTCCCAAAACATGTTATCTTTACACCAACCGGCTGCTCA 161
Qy      214 TATCCGCTGCTTATTTCTTCATGAGCTTTTATCTTGCAACTACTTCTACAGCTT 273
Db      162 TACCTGTATATTTGTTCTGCAATGATTTTCCCTTGCAATAGCTACTGAGCTC 221
Qy      274 CTTAACACATCGCTTGCAATGTTAATCTTGTAGCCCAAGTTGTGCAATTATTTG 333
Db      222 CTAGGCCACATAGCTTCAATGATGATTAATTTGCTCCTCAGCTGTGTGAGTGT 281
Qy      334 CCGCC-----GGAGGGCAAGTGAAGTGAAGTCTGGAATGTGTAATCTG--- 384
Db      282 GGGCTATGTTGGAACCTGCTGATGAATTAATTTTCAAGGAAAGTTGGATTTGCTA 341
Qy      385 GCATCGGAAAACTCAAGCTCACTTACCAACTTCGTTAAATGTTATGAAAAATACACC 444
Db      342 GCGAGAGAGGGGCTTCAACTCTGCTTCCAGAGAAATTTGAAGCAAAATTTGATTAATG 401
Qy      445 TCACTGCTGGGCAAGCCGCGGTGGGAAAAAGGCGTTTGGGCTTGAAGCCATGCC 504
Db      402 GTTTTATCAGGTCAACAAAGGGTGGCAAACTGTATTTGCTGTGGCACTTGTATGCT 461
Qy      505 GCAACATTTAGACCATCATCAAGTTTTCAGCTCTAATAGAAATTTGATTCAGTCGACGA 564
Db      462 AAAA-----CTAACTCAAGTTTTCAGCACTTAGAGCAATGACCTGTGGCTGGC 512
Qy      565 ACTAACAAATATCATTAAGACGATCCGATATCTTAACGTATTAACCGGAAATCTTTGAG 624
Db      513 CCAATGAATCTTGTGGAACATTTCTCTATTTCTCACTGSCAATGTCCCAATCTTCAAT 572

```

```

Qy      625 CTGACATATACCGGTTGAGTGTGGGAAACCGGACTCGGACC---GAATGGAACACATGT 681
Db      573 TTGAACATACCAATGTTGTAATTTGGCACTGGGCTAGGCCOAGAGAGGCTAATTTT 632
Qy      682 ATGCACATTCGCGACCAAGGACTTAACCATGAGAGTTTTCAAAGATGTAAGCG 741
Db      633 ATTCAACATGTGTCTCTGATGAGGTGACCAATTAAGAGATTTTCAATTAAGTGAACCC 692
Qy      742 ACGAAACCCATTTCTGCTGCGGATTTACGACATATGATATGTTGACGATGATTTG 801
Db      693 CTTGTCAATATTTTGTGCAACTGATATGTCATGACATGATTTGATGATGTACA 752
Qy      802 CCGGTTTTGTGGGT---TTATGCGCGTTGATGTAAGATGGGCAAGAAAAAG 858
Db      753 CTTGCTTAATTTGGGCTCAATATTTGTCAAAATTTGATATGCAAGATGGGAAGGTCCTAAG 812
Qy      859 TCTGAGATGAGAGCTTTGATGAGTGAATTTGTGCGTTTCTCAAGTATAGTTTGTG 918
Db      813 GACTGATGAGAAAGAACCGGAGAGGAGGTGTGTGCTTTCTTAAGGGCACAGTTGAAT 872
Qy      919 GGTGAAAAACGGAGATTGATGATTGATGAGAGATCTTCGTTTCTCCGCGCAAGCTT 978
Db      873 GCGCTATGGAAGATTTTAATGCTGTTTGTGGCAATCTAATCTTGTCTTAACTG 932
Qy      979 GAT 981
Db      933 GAT 935

```

RESULT 7

```

US-10-381-123-17
; Sequence 17, Application US/10381123
; Publication No. US20050081263A1
; GENERAL INFORMATION:
; APPLICANT: E.I. du Pont de Nemours and Company
; TITLE OF INVENTION: Chlorophyllases
; FILE REFERENCE: B01477 PCT
; CURRENT APPLICATION NUMBER: US/10/381,123
; CURRENT FILING DATE: 2003-03-21
; PRIOR APPLICATION NUMBER: PCT/US01/31059
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: 60/238,161
; PRIOR FILING DATE: 2000-10-05
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 17
; LENGTH: 1242
; TYPE: DNA
; ORGANISM: Trifolium aestivum
US-10-381-123-17

```

```

Query Match      16.8%; Score 199; DB 9; Length 1242;
Best Local Similarity 55.8%; Pred. No. 4.5e-49;
Matches 476; Conservative 0; Mismatches 350; Indels 27; Gaps 4;

```

```

Qy      137 CCGTGAATCCGTTGAAAAATGATTCAACAGCACCGCCAAACCGGTGAGATCACTGTCT 196
Db      153 CGATTCAGGTGATGATGAAATGGGGCCGACACACAGATCCGGTGTGATGTGGCAC 212
Qy      197 CAACAGTCCGCGGAATTTATCCGCTGTTTATTTCTTCAATGAGCTTTTATCTTGCAACT 256
Db      213 CCAAGATGAGAGAACCTACCCCGTGCATGTCTTTGACAGGCTTCTTCTCCATTAACC 272
Qy      257 ACTTCTACTGTGAGTTTCTTAAACCATGCGCTTGAGATGATTAATTTTGTAGCCCAAC 316
Db      273 ACTTCTACGAACCTTCTTCCGCACTGCGATCCACGCGCTTCAATTTGTGCGCCCC 332
Qy      317 AGTTGTG---CAATTAATTTCCGCGCGGAGGCGCAAGTGAAGTGAAGTGTGAAAGTG 373
Db      333 AGTTGAGCATCATATCATACCTTGGGGTGAAGGACAGACATCGCCGGGCGACCAAG 392
Qy      374 TGATTAACGTGGGATCGGAAAAACCTCAAGCTCCTACCAACTTGGTAAATGCTAATG 433

```

```
Db 393 TGGCAGACTGGCTCCCGACGGCTCTCCGCTCGTGTCTGCCCCAAGCGCTCGACGGCAGC 452
| | | | |
Qy 434 GAAATATACACTCTACTCTGTGGGCGACAGCCGGGTGGGAAAACGGCTTTGGGTGGC 493
| | | | |
Db 453 TCTCGAAGCTCGCTTGGCCGGCCACAGCCGAGGAGCCACACGGCTTTCTCCCTGGCCT 512
| | | | |
Qy 454 TAGGCAATGCCGCAATTAGCCATTCATCCATTTTCACTTAAATAGAAATTTGATC 553
| | | | |
Db 513 TGGGCGACGCC-----AAGACCCAGCTAACCTTCTCGCGCTCATTCGACTCGACC 563
| | | | |
Qy 554 CAGTCGACGAACTAACAAATACATTAGAACCGATCCGATCTTAAAGTAAACCGG 613
| | | | |
Db 554 CCGTCGCGGCGACGGGGAAAGCTCTCCAGCTCCAGCCAGATCTCTACATCGAGCCGT 623
| | | | |
Qy 614 AATCTTTGAGCTGACATACCGGTTGACAGTGTGGAAACGGGACTCGGACCGAAGTGA 673
| | | | |
Db 624 CCTCTTGGGCAATGCGAGATGCGGTGTGTATCGGACCGGGCTCGGCGAGAGAAAG 683
| | | | |
Qy 674 ACAAGTGATG---CCACCATGCGGACCAACGGAATTAAACATAGAGATTTTCAAA 730
| | | | |
Db 684 AGAATATATTTCTTCCCTCCCTGCGACCCAGAGAGTGAACACCGGAGTTCTACGCG 743
| | | | |
Qy 721 AGTGAAGGCGACGAAAGCCATTTCTGTGCTGCGGATTAACGATATGATATGTGG 790
| | | | |
Db 744 AGTCAGGCGCGCTCTGCTACTACTTTGTGACCAAGATCTACGCGGATCTGACATGCTGG 803
| | | | |
Qy 791 ACGATGATTTGCGCGGTTTGTGGGTTTATGCGCGTTGTATGTGAAGATGGGCAAA 850
| | | | |
Db 804 ACGAGAGCGCCCGCAAGTTCTAT-----CACCTGCTGTGCAAGATGGGAAAG 851
| | | | |
Qy 851 GAAAAAGCTGTGATGAGAGAGCTTTGTAGTGAATGTGTGCTGCTTCTCAAGTAA 910
| | | | |
Db 852 GGTGCAAGGCGAAGATGCGAGGTGCTGTGCTGAGATCATGTGCGCATTTCTTAATGCTG 911
| | | | |
Qy 911 GTTTGTGGGTGAAAAAGCGGAGATTCGATTTGTGAGAGATCTTCCGTTTCTCCGG 970
| | | | |
Db 912 CTTTGGGTGAGAAAGATGAGATCTTTGAGGCAATCTGAGAGACCGGGGCTTGACCCA 971
| | | | |
Qy 971 CCAAGCTTGATCC 983
| | | | |
Db 972 CCACGCTTGATCC 984
| | | | |

RESULT 8
US-10-381-123-9
; Sequence 9, Application US/10381123
; Publication No. US20050081263A1
; GENERAL INFORMATION:
; APPLICANT: E.I. du Pont de Nemours and Company
; TITLE OF INVENTION: Chlorophyllases
; FILE REFERENCE: B81477 PCT
; CURRENT APPLICATION NUMBER: US/10/381,123
; CURRENT FILING DATE: 2003-03-21
; PRIOR APPLICATION NUMBER: PCT/US01/31059
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: 60/238,161
; PRIOR FILING DATE: 2000-10-05
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 9
; LENGTH: 1104
; TYPE: DNA
; ORGANISM: Glycine max
US-10-381-123-9

Query Match 16.4%; Score 194.4; DB 9; Length 1104;
Best Local Similarity 56.3%; Pred. No. 1e-47;
Matches 473; Conservative 0; Mismatches 346; Indels 21; Gaps 5;

Qy 154 AATGATTCAGACGACCGCAAAACCGGTGAGAAATCACTGTCCAAAGTCCGGAACT 213
| | | | |
Db 113 AGTGCTTCTCTCACTCAACCAACCATTTGTAATTTTACACCAAGTGTCTGCGCGCA 172
| | | | |
```

```
Qy 214 TATCCGTCGCTTTTATTTCTTCATAGCTTTTATCTTGGCAATCTTCTACTGACGTT 273
| | | | |
Db 173 TACCTGTAAATATTTGTCATAGCTTTTCAATTCGCAATTTCTACTCTCAAGCTC 232
| | | | |
Qy 274 CTTAACCAATCGCTTGCATAGTTTAACTTTGTAGCCCC--ACAGTTGCAAAATTA 330
| | | | |
Db 233 CTAAACCACATAGTCTCAATGATTTATATATGTTGCTCTCACTGTTTTCATAGGG 292
| | | | |
Qy 331 TTGCGCGCGGAGGCGAAGTGAAGTGAAGTGAAGTGTGATTAATCTGGGCAATCG 390
| | | | |
Db 293 CTTCCTATGTATGAGACCACTGAAAGTGAATATGCAAGAAAGTTGGGATTTGATAGCT 352
| | | | |
Qy 391 GAAATCCTAAAGTCACTACCACTTCCGTTAAATGTAATGAAATATACCTCACTC 450
| | | | |
Db 353 GAGAGGCTTCAACATTTGCTTCCAGAGACGTTGAGCAATTTGGACAACTGGTTCTA 412
| | | | |
Qy 451 GTGGGCGACAGCCCGGTGGGAAACGCGCTTTGCGGTGCTAGGCCATGCGGCAACA 510
| | | | |
Db 413 TCAAGTCAAGTATAGGGGTGGGAAACGTATATTTGCTGTGCTCTTGTGCTATGC----- 465
| | | | |
Qy 511 TTGAACCATTCATACAGTTTTCAGCTCTTAATAGAAATGATTCAGTGCAGAACTAAC 570
| | | | |
Db 466 --AAAACTAATCTCAAGTTTTCAGCACTTGTAGGCAATAGACCTGTGGCTGGCAATCT 523
| | | | |
Qy 571 AAATACATTTGAACCGAATCCGCAATCTTAAAGTAAACCGGAATCTTGGAGCTGAC 630
| | | | |
Db 524 AAATATTTGAGAACGCTCTCATATTTCTCATGCGCAAGCCAGGCTCTTGAATTTGAA 583
| | | | |
Qy 631 ATACCGGTTGAGTGTGGAAACCGGACTCGGACC--GAAGTGAACAAAGTATGCA 687
| | | | |
Db 584 ATGCAATTTGAAGTAATTTGGACATGGAATGGGCCAGAGAAAGTTATGTTGAATCTCA 643
| | | | |
Qy 688 CCAATGCGCACCAACGACTTAACCATGAGAGATTTTCAAAAGTGTGAAGCGACGAAA 747
| | | | |
Db 644 CCGGTGTCTGTGATGGGGTGAACATTAAGAGATTTCTCAACGAGTGAACCCCTTGT 703
| | | | |
Qy 748 GCCATTTTCGAGCTGCGGATTAACGACATATGATATGTTGACGATTTTCCGCGCT 807
| | | | |
Db 704 GCTAAATTTTGTGAGCAAAAGTATGCTACATGACATGTTGAATGATGACACACAGG 763
| | | | |
Qy 808 TTTGTTGGG--TTTATGCGCGGTTGTATGTATGTAAGATGGGCAAA--GAAAAAGCTCT 861
| | | | |
Db 764 CTAATTTGGACATTTGATGTCAAAGTATGTAAGATGAGAGACGACGGGTCTTAGGGAC 823
| | | | |
Qy 862 GAGATGAGAGCTTTGATGATGGAATGTGTGCTGCTTCTCAAGTATAGTTTGTGGGT 921
| | | | |
Db 824 TTGATGAGAAAGAACCACTGAGGGGTGTGTGCTTCTTGAAGGCACAATTTGAATGAC 883
| | | | |
Qy 922 GAAAAAGCGGAGATTCGATTTGATGTAAGAGATCTTCCGTTTCCGCGCAAGCTTAT 981
| | | | |
Db 884 CTATGAGAAAGATTTTGTATGCTATTTTAAAGAACCTTAATCTTGCTCCCATGAACTGAT 943
| | | | |

RESULT 9
US-10-381-123-11
; Sequence 11, Application US/10381123
; Publication No. US20050081263A1
; GENERAL INFORMATION:
; APPLICANT: E.I. du Pont de Nemours and Company
; TITLE OF INVENTION: Chlorophyllases
; FILE REFERENCE: B81477 PCT
; CURRENT APPLICATION NUMBER: US/10/381,123
; CURRENT FILING DATE: 2003-03-21
; PRIOR APPLICATION NUMBER: PCT/US01/31059
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: 60/238,161
; PRIOR FILING DATE: 2000-10-05
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 11
; LENGTH: 1125
; TYPE: DNA
```


Qy 918 GGTGAAAAGCGAGATTCGATTGATGTGAGAGATCCCTTCCTTCTCCGGCCAACT 977
Db 927 TGGCCTATGGAAGATTAAATGCTGTTTGCGAGATCTTAATCTGCTCTACTAACT 986
Qy 978 TGAT 981
Db 987 GGAT 990

RESULT 11
US-10-381-123-13
; Sequence 13, Application US/10381123
; Publication No. US20050081263A1
; GENERAL INFORMATION:
; APPLICANT: E.I. du Pont de Nemours and Company
; TITLE OF INVENTION: Chlorophyllases
; FILE REFERENCE: BB1477 PCT
; CURRENT APPLICATION NUMBER: US/10/381,123
; PRIOR FILING DATE: 2003-03-21
; PRIOR APPLICATION NUMBER: PCT/US01/31059
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: 60/238,161
; PRIOR FILING DATE: 2000-10-05
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO: 13
; LENGTH: 1444
; TYPE: DNA
; ORGANISM: Gesneriana
US-10-381-123-13

Query Match 15.1%; Score 178.8; DB 9; Length 1444;
Best Local Similarity 54.1%; Pred. No. 6,8e-43;
Matches 438; Conservative 0; Mismatches 357; Indels 15; Gaps 3;

Qy 156 TGATTCACAGCAGCCGCAAAACCGTGAGATCACTGTCACAGTCGCCGGAACCTTA 215
Db 195 TGTCTCCGACCACTCCGAGAGCCACCACTGATCCCACTCATTTGAGAAAGGAGTA 254
Qy 216 TCCGTCCTTTTATCTTCATGCTTTTATCTTCGCACTACTTCTGACGTTCT 275
Db 255 CCCAAGCTTCTACTCTTCATGATCATGCTTCACAACTTCTACTGAGCTTAT 314
Qy 276 TAACACATCGCTTCGAGATGATCATCTGTGAGCCCAAGTGGCAATTAATGCC 335
Db 315 CGAGCATCGATCCATGCTTATGTTGTTGATCTCACTTAATCTTGATGCTAC 374
Qy 336 GCCGAGAGGAGAGTGAAGTGAAGTGAAGTGAAGTGAATGAATGAGGATCGAAAA 395
Db 375 ATGTATATGTAAGATGAGATGATGATGATGATGATGATGATGATGATGATGATG 434
Qy 396 CCTCAAGCTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACT 455
Db 435 ACTGCAATGTTCTCCCAACAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 494
Qy 456 CCACAGCGCGGTGGGAAAAAGCGGTTGCGGTGCGGTGCGGTGCGGTGCGGTGCG 515
Db 495 ACATGCGCGGTGGGAAAAAGCGGTTGCGGTGCGGTGCGGTGCGGTGCGGTGCG 545
Qy 516 CCCATCATCACTTTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACT 575
Db 546 GACTCATATTAAGCTTCTCAGGCTGATGAGGATGATGATGATGATGATGATGATG 605
Qy 576 CATTGAACCGATCCGATATCTTAAGTGAATGATGATGATGATGATGATGATGAT 635
Db 606 ACACCAAACTAATCTCTGATCAATTAATTAATTAATTAATTAATTAATTAATTA 665
Qy 636 GATTGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 695
Db 666 GTCATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 722
Qy 696 ACCAAGCATTAACATGAGAGTGTAAAGAGTGTAAAGAGTGTAAAGAGTGTAAAG 755

Db 723 CCCAAGAGATTAATACCAAGATTTTACAGATGTTCTTCCCGCTTGCATTT 782
Qy 756 CGTGCCTCGGATTAACGACATATGATATGATATGATATGATATGATATGATATG 815
Db 783 TGTGCTAAGAGATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 842
Qy 816 GTTATGCGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 875
Db 843 TATTGCTCATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 899
Qy 876 TGTAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 935
Db 900 TGTAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 959
Qy 936 TCGATTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 965
Db 960 GCGAGCTGTAAGATTAACCCCGAGATTTT 989

RESULT 12
US-10-381-123-3
; Sequence 3, Application US/10381123
; Publication No. US20050081263A1
; GENERAL INFORMATION:
; APPLICANT: E.I. du Pont de Nemours and Company
; TITLE OF INVENTION: Chlorophyllases
; FILE REFERENCE: BB1477 PCT
; CURRENT APPLICATION NUMBER: US/10/381,123
; PRIOR FILING DATE: 2003-03-21
; PRIOR APPLICATION NUMBER: PCT/US01/31059
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: 60/238,161
; PRIOR FILING DATE: 2000-10-05
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO: 3
; LENGTH: 1216
; TYPE: DNA
; ORGANISM: Vitis sp.
US-10-381-123-3

Query Match 14.3%; Score 169.8; DB 9; Length 1216;
Best Local Similarity 53.2%; Pred. No. 3,4e-40;
Matches 435; Conservative 0; Mismatches 367; Indels 15; Gaps 3;

Qy 167 CACGCCAAAACCGTGAAGATCACTGTCCACAGTCGCCGGAATTAATCCGCTTT 226
Db 112 CCCCTCCACTGCACTTGAATGCTAGGCCCTCTGAAGCAGGAGTCCCGCTGCTGC 171
Qy 227 TATTCCTCAATGCTTTTATCTTCGCACTACTCTCTGACGTTCTTAACACATGCG 286
Db 172 TCTCTCTCAATGCTTTTATCTTCGCACTACTCTCTGACGTTCTTAACACATGCG 231
Qy 287 CTTCGATGTTACATCTTGTAGCCGCCAGTGTGCAATTAATGTCGCCGGAAGGC 346
Db 232 CCTCTCAATGTTATGTTCTTGTAGCCGCCAGTGTGCAATTAATGTCGCCGGAAGGC 291
Qy 347 AATGGAAGTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 406
Db 292 GCGAAGGATCAATCCGACAGCTGTTTAACTTAATGTTTAACTTAATGTTTAACT 351
Qy 407 ACCAACAATGCGTAAATGCTAATGGAATTAATGATGATGATGATGATGATGATGAT 466
Db 352 TACTTCTCCCATGTTCCGCAATTTAAAGCAATTAAGCAATTAAGCAATTAAGCA 411
Qy 467 GTGGAAGAGCGGCTTTGCGGTTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCG 526
Db 412 GAGCAAAAGCTTTTGTCTGACACTGAG-----AAAGCATTCACCTTCTCTGA 462
Qy 527 CGTTTCACTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 586
Db 463 AATTTCAAGCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 522

QY 587 ATCCGATATCTTAACGATTAACCGGAATCTTTGAGCTGACATACCGGTTGCACTGG 646
 Db 523 CTCACCGGTACTACCTATGTTCTCTCATTTGATGATGACATGCGATGATGTA 582
 QY 647 TGGGAACCGGACTCGG---ACCGAAGTGAACAACGTATGCAACATGCGCACCAACGG 703
 Db 583 TTGGTTCGGGTTTGGGTGAAGTGAAGAAAGAACCTCTTCTCTCTCTTGTGCCCCAAGG 642
 QY 704 ACTTAACCATGAGAGATTTTACAAAGATGTAAGCGACGAAAGCCATTTCTGTGCTG 763
 Db 643 GCGTAACCATGAGAGACTTTCTTTAAAGAAATGCCGTAACAGCTGTATTTCTTGCA 702
 QY 764 CGGATTAACGATGATGATATGTTGAACGATGATTTGCCGTTTGTGGTTATG 823
 Db 703 AGGACTATGAGCATCTTGAACATGCTAGACGATGAGACTAATGAAATTAAGAGAAAGCTA 762
 QY 824 CCGGTTGATGATGATGATGATGAGGCAAAAGAAAGTCTGATGATGAGAGCTTTGTAGGTG 883
 Db 763 CACATTTGTTGTGTAAGAAATGGAAGTGAAGAAC---CATAGAGAGTTTGTGAG 819
 QY 884 GAATTTGTGTTGCGTTTCTCAAGATATGTTTGTGGGTGAAGAAAGCGAGATTCGATTA 943
 Db 820 GCATTTGATGATGATTAATGAAAGCTTATTTGAAAGGAGATTAACAGCAGTCTAATCTCA 879
 QY 944 TTGTGAAGATCCTTCCGTTTCTCCGCGCAAGCTTGA 980
 Db 880 TTAGAGATGGGATGCTAGTACAGCACAGTGGAGCTTCA 916

RESULT 13

US-10-634-548-19
 ; Sequence 19, Application US/10634548
 ; Publication No. US20040045051A1
 ; GENERAL INFORMATION:
 ; APPLICANT: No. US20040045051A1, Susan R
 ; APPLICANT: Lincoln, Kim
 ; APPLICANT: Abbad, Mark Scott
 ; APPLICANT: Eliezer, Robert
 ; APPLICANT: Hartshuber, Karen Kindie
 ; APPLICANT: Hirschberg, Joseph
 ; APPLICANT: Karunanandaa, Balasubramanian
 ; APPLICANT: Moshiri, Farhad
 ; APPLICANT: Stein, Joshua C.
 ; APPLICANT: Valentin, Henry E.
 ; APPLICANT: Venkatesh, Tyamagondlu V.
 ; TITLE OF INVENTION: Tocopherol biosynthesis related genes and used thereof
 ; FILE REFERENCE: Ren-01-125
 ; CURRENT APPLICATION NUMBER: US/10/634, 548
 ; CURRENT FILING DATE: 2003-08-05
 ; PRIOR APPLICATION NUMBER: us 60/400, 669
 ; PRIOR FILING DATE: 2002-08-05
 ; NUMBER OF SEQ ID NOS: 79
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 19
 ; LENGTH: 1135
 ; TYPE: DNA
 ; ORGANISM: Arabidopsis thaliana
 ; US-10-634-548-19

Query Match 13.8%; Score 163.4; DB 7; Length 1135;
 Best Local Similarity 54.3%; Pred. No. 2.9e-38;
 Matches 420; Conservative 0; Mismatches 341; Indels 12; Gaps 4;

QY 157 GATTCAACAGACCGCAAAACCGGTAGATCACTGTCCACAGTGGCGGAATTTAT 216
 Db 151 GCTTCAACGCTCTCGCCAAAGACACTGTGTGCTACGCCGTGAGAGAAAGATTTAT 210
 QY 217 CCGCTGCTTTATTTCTTCATGCTTTTATTTTTCGACATTAATCTTTCAGCTTCTT 276
 Db 211 CCGGTGATGATGCTTCTTCATGCTTTTATTTTTCGACATTAATCTTTCAGCTTATG 270
 QY 277 AACACATCGCTTGACATGTTATCTTGTAGCCCAACAGTTTGCATAATTAATTCGCG 336

Db 271 TTGATGATCTCTTCTCACTGAGCTTATCTATGCTCTCTAGTTATATATATATGCCGGA 330
 QY 337 CCGGAGAGGCAAGTGAAGTGAAGCATGCTGGAAGTGTATTAACGTGGATCGGAAGAAC 396
 Db 331 CCAGACCATATGATGATGATTAATCAACCGCGAGATTAATGATTTGTTATGATGGA 390
 QY 397 CTCAAGCTCACCTTACCACTTGGTAAATGCTAATGGAATAATACCTTCATCTGTGGGC 456
 Db 391 CTTAATCATCTTCTTCCAGCGCAAGTAACACCAACTATTCAAATTTGGCTCTCCGGC 450
 QY 457 CACAGCCGCGGTGGGAAACCGGCTTGGCGGTGGCGGTAGCCATGCCGCAATTAAGAC 516
 Db 451 CATAGCCGCGGTGGCAAAACCGGTTGGCGGTGGCGCTTAAAGAAATTTG---GGTACTCC 507
 QY 517 CCAATCCATCACTTTTCACTTAAATGAAATGATTCAGATCCAGAACTTAACAAATAC 576
 Db 508 TCGAATCTAAGATCTGACATTTGATGATGATGATGATGATGATGATGATGATGATGATG 567
 QY 577 ATTGAACCGATCCGATATCTTAAAGATTAACCGGAATCTTTGAGCTGACATACCG 636
 Db 568 AAACAAACCCCTCTCCGCTGCTTACCTTCAAACTCATTTGACCTTAAGCAAAACG 627
 QY 637 GTTGCA---GTGATGGGAAACCGGATCCG---ACCGAAGTGAACAACGATGCCACCA 690
 Db 628 CTATATCTTGTATGCTGCTTGGGCTTGGTGAACCGCTCGAACCCATTAATTTCCACCG 687
 QY 691 TGGCACAACACGCACTTAAACATGAGAGATTTTAAACAGAGTGAAGCGAGAAAGCC 750
 Db 688 TGTGCACTCCCGGAGATGATCAACCGAGATTTCTTTCGGGAAATGTCAAGTCCAGCATG 747
 QY 751 CATTTGCTGCTCGGATTAACGACATATGATATGTTGACGATGATTTGCCCGTTTT 810
 Db 748 CATTTGCTGCGAAGATTAATGAGCATTTGGAACATGCTGATGATATCAAAAGGATTT 807
 QY 811 GTTGGGTTTATGCGCGGTTGTATGATGATGATGATGATGATGATGATGATGATGATG 870
 Db 808 AGAGGGAAGAGTTCTTATTTGTTTGTGTAAGAAATGGA---AGAGGAGAGACCAATGAGG 864
 QY 871 AGCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 923
 Db 865 AGATTCGTTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 917

RESULT 14

US-10-424-599-34940
 ; Sequence 34940, Application US/10424599
 ; Publication No. US20040031072A1
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa Thomas J
 ; APPLICANT: Kovalic David K
 ; APPLICANT: Zhou Yihua
 ; APPLICANT: Cao Yongwei
 ; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
 ; FILE REFERENCE: 38-21(53223)B
 ; CURRENT APPLICATION NUMBER: US/10/424, 599
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 285684
 ; SEQ ID NO 34940
 ; LENGTH: 1068
 ; TYPE: DNA
 ; ORGANISM: Glycine max
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: PAT_MRT3847_131553C.1
 ; US-10-424-599-34940

Query Match 11.4%; Score 135.6; DB 7; Length 1068;
 Best Local Similarity 53.2%; Pred. No. 7.9e-30;
 Matches 447; Conservative 0; Mismatches 354; Indels 39; Gaps 6;

QY 154 AATGATTCAACAGCACCGCAAAACCGGTAGATCACTGTGTCACAGTCCGCGGAATCT 213

```
Db 91 AGTGTCTCTTCTCACTCCAAAACCATGTGTAATTTTACACAACTGTTCTGGCGCA 150
Qy 214 TATCCCGCTTTTATTTCTCCATGCGCTTTATCTTCGCACTACTTACTGACCTT 273
Db 151 TACCTGTAAATATGTTGTGTCATGCGCTTTTCATTCGCAATTTCTACTCAAAAGCTC 210
Qy 274 CTTAACCAATCGCTCCCATGTTACATTTCTTGTAGCCCAAGTTGTGCAATTTATG 333
Db 211 CTAGCCCAATAGTTCACATGACAGATTGAGACAAAGATTAATGATTCATGAGAAC 270
Qy 334 CCGCCGCG--AGGCCAAGTGAAGTGAAGATGCTGGAAGTGTATTAACGTGGCATCG 390
Db 271 TATGTGGCTAGAGACTTTAGATAGTAATAGAGAAAAGTTGCCGATTTGATAGCT 330
Qy 391 GAAAACCTCAAGCTCACCTACCACTTCGTAATGTGAATGAAAATATACCTCATC 450
Db 331 GAGGAGCTTCAACATTTGCTCCAGAGAACGTTGAAGCAATTTGAGCAAACTGGTTTA 390
Qy 451 GTGGCCACAGCCGCTGGGAAAACGGCGTTTGGCTGAGCCATGCGCAACA 510
Db 391 TCAGGTCAAGTAGGGGTGGGAAAACGTATTTGCTGTGCTTGTGTCATGC----- 443
Qy 511 TTAGACCATTCATCAGCTTTTCACTCTTAATGGAATTTATCAGTGCAGAACTAAC 570
Db 444 -AAAACTAATCTCAAGTTTTCAG-----CACTGTAGGCACATCT 483
Qy 571 AAATCATTAAGACCGATCCGATATCTTAACGTATAACCGGAATCTTTGAGCTGAC 630
Db 484 AAATATTGTAGAACGTCCTCATATTTCTCACTGCAAGCCAGCTCTTGTATTTGAA 543
Qy 631 ATACCGGTTGCACTGCTGGGAAACCGGACTCGGAC--GAAGTGAACAAGTATGCCA 687
Db 544 ATGCAAGTTGAAGTATTTGGCACTGGATTGGCCCAAGAGACTTAATTTGTATCTCA 603
Qy 688 CCATCGCAACCAAGCACTTAACCATAGAGATTTCAAAAGAGTGAAGGCGACGAA 747
Db 604 CCGTGTCTCTGATGGGGTGAACATAAGAGTTCTTCAACGAGTGAACCCCTTGT 663
Qy 748 GCCCATTCGTGGCTGGGATTAAGGACATATGATGATGAGATGATTTGGCGGCT 807
Db 664 GCTAATTTGTGTAAGAAAGTATGTCACATGACATGATGATGATGACACACAGG 723
Qy 808 TTTGTTGGG---TTATGCGCGGTGTATGTTAAGAAAGGCAAA---GAAAAAGTCT 861
Db 724 CTAAATTTGGGACATTTGTGTCAAAAGTATGTTAAGATGGGACACGCGTCTAGGAC 783
Qy 862 GAGATGAGAGCTTTGTAGTGAATTTGTGTTGCTTCTCAAGTATATTTGTGGGT 921
Db 784 TTGATGAGAAGGACCACTGAGGGGTTGTGTCTTCTTGAAGGCAAAATTTGATAC 843
Qy 922 GAAAAAGGAGATTCGATGATGTTGTAAGATCTTCGCTTCTCGGCGCAAGCTGAT 981
Db 844 CTATGGAAGATTTGATGCTATTTTAAAGGACCTTAATCTTGCTCCACATGAGTCAT 903

RESULT 15
US-10-381-123-25
; Sequence 25, Application US/10381123
; Publication No. US20050081263A1
; GENERAL INFORMATION:
; APPLICANT: E.I. du Pont de Nemours and Company
; TITLE OF INVENTION: Chlorophyllases
; FILE REFERENCE: B01477 PCT
; CURRENT APPLICATION NUMBER: US/10/381,123
; PRIORITY FILING DATE: 2003-03-21
; PRIOR APPLICATION NUMBER: PCT/US01/31059
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: 60/238,161
; PRIOR FILING DATE: 2000-10-05
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 25
; LENGTH: 1156
```

```
; TYPE: DNA
; ORGANISM: Trilecium aestivum
US-10-381-123-25
```

```
Query Match 9.3%; Score 110.8; DB 9; Length 1156;
Best Local Similarity 51.1%; Pred. No. 2.98-22;
Matches 436; Conservative 0; Mismatches 382; Indels 36; Gaps 6;
```

```
Qy 169 CCGGCAAAACCGGAGAAATACCTGTCCAAAGTCCCGGAATTTATCCGTGTTTA 228
Db 141 CTGGCCAGCCGTTGATGAGTGTGGCGCCCAAGAGACCGCGGTACCCGTATGCTC 200
Qy 229 TTCTTCATGCTTTTATCTTCGCACTACTTCTACTCTGACGTTCTTAAACCATGCT 288
Db 201 TTCTTCGACGCTTGAACATGCTCAACAGCTGTGACAGAGCTCTCTACACAGTGGC 260
Qy 289 TCGCATGTTACATTTCTTGTAGCCCAAGTTGTGCAATTTATTTGCCGCGGAGGCA 348
Db 261 TCCCATGTTTCAATCGCGCTGACACAGCTTACTGATGTGTCCGAGCCGATGCG 320
Qy 349 GTGAAAGTGAAGATGCTGGAAGTGTATAACTGGGCAATCGGAAACCTCAA--AGCTC 406
Db 321 GACGACATAGACGCAAAAGCGAATCAACAATGCGCTTGACATGACAAAGGCTC 380
Qy 407 ACTTACCAACTTCGCTAATGCTAATGAAAAT-----ACACC 444
Db 381 GCCACGTCCTCAAGACAGTGTCAAACTTGAAGATGCAAGCTGACCTGTCCAGCTG 440
Qy 445 TCACCTGTGGCCACAGCCGCGTGGGAAAACGGCGTTTGGCGTTGCGCTAGGCCATGC 503
Db 441 GCCCTACCGGCAATACGAGGCGGCGACAGCGGCTTTCGCGCTGCGCTGGGACTAGGG 500
Qy 504 --GSCAATTAAGCCCATCATCAGCTTTTCACTCAATTAAGAAATTTGATTCAGTGC 561
Db 501 GACCCAAAGACCAAGCTGAGCTCAAGTCTTCGCTCTCATCGGCTGACCCGTTGCC 560
Qy 562 GAACTAACAATCATTAAGACCGATCCGATATCTTAAAGTATAACCGGAATCTTTC 621
Db 561 GGGGTTTCCAGAGCCCAAGTTGAGAGCCCAAGGTGCTCATTTTGAACATGATCTC 620
Qy 622 GAGCTGACATACCGTTGCACTGTGGGAAACCGGACTCGGACCGAAGTGAACACGTG 681
Db 621 GACCTGGGAGTCCGCGTGTGTCATGAGGACTGGGCTGGGCTGCCAA---GCACATGCGC 677
Qy 682 ATGCAACATGCGCAACAAGGACTTAACCATAGAGAGTTTTCAAAAGGTGAAGGCG 741
Db 678 GGAATTCATAGCGCCCGGTGGGCGTGAACACCGGAATTTCAAGAGAGTGGCGCG 737
Qy 742 ACGAAACCCATTTCTGTGCTGCGGATTAAGACATATGATATGTTGACATGATTTG 801
Db 738 CTTGCTACACACCTCGTGTGTCAAGAGCTACGGGCACTTCGACATGCTGAGACATGTG 797
Qy 802 CCGCGTTTGTGGGTTATGCGCGTTGTATGTTAAGAAAGGCAAAAGAAAGTCT 861
Db 798 CCC---TATATCATCAACAACATGATGATGAGAAACCAACGACACCAAGATCTT 854
Qy 862 GAGATGAGAGCTTTGATGAGAAATTTGTTGCTTCAAGTATATGTTATGTTGGGT 921
Db 855 GCTAGAGAGACATAGAGAGACCA---TGTATCTTCTCTCAGAGCTTAATTTGCAATC 911
Qy 922 GAAAAAGGAGATTCGATGATGTTGTAAGATCTTCGCTTCTCGGCGCAAGCTTAT 981
Db 912 GATTTGTGATCTCATGCGCATATATATCTGAGATCGGCGCAGCGGCTCTGAC 971
Qy 982 CTTTCACTGAGTT 995
Db 972 CAAGTTGATGAGTT 985
```

```
Search completed: March 20, 2006, 14:42:17
Job time : 1105.06 secs
```

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 20, 2006, 14:07:47 ; Search time 475.098 Seconds

(without alignments)
5830.519 Million cell updates/sec

Title: US-10-634-548-18

Perfect score: 1188
Sequence: 1 gatccataaattctcaacac.....ccccaaaaaaaaaaaaa 1188

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 8023312 seqs, 1165852854 residues

Total number of hits satisfying chosen parameters: 16046624

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA New:
1: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB_seq.*
2: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB_seq.*
3: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB_seq.*
4: /cgn2_6/ptodata/2/pubpna/PCF_NEW_PUB_seq.*
5: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB_seq.*
6: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB_seq.*
7: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB_seq.*
8: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB_seq.*
9: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB_seq.*
10: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB_seq.*
11: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB_seq.*
12: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB_seq.*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB_seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	39.6	3.3	542	US-09-925-065A-175209	Sequence 175209,
2	38.6	3.2	958	US-09-925-065A-683561	Sequence 683561,
3	37.6	3.2	534	US-09-925-065A-102394	Sequence 102394,
4	37.6	3.2	543	US-09-925-065A-849244	Sequence 849244,
5	37.4	3.1	605	US-09-925-065A-902549	Sequence 902549,
6	37.2	3.1	578	US-09-925-065A-358029	Sequence 358029,
7	37	3.1	590	US-09-925-065A-106455	Sequence 106455,
8	37	3.1	594	US-09-925-065A-417521	Sequence 417521,
9	37	3.1	599	US-09-925-065A-106454	Sequence 106454,
10	37	3.1	893	US-09-925-065A-283413	Sequence 283413,
11	37	3.1	566	US-10-240-708-29	Sequence 29, Appl1
12	36.4	3.1	1691140	US-11-091-018-1	Sequence 1, Appl1
13	36	3.0	617	US-09-925-065A-881350	Sequence 881350,
14	36	3.0	1462	US-10-750-185-46341	Sequence 46341, A
15	36	3.0	1462	US-10-750-185-46341	Sequence 46341, A
16	36	3.0	1830	US-10-750-185-38673	Sequence 38673, A
17	36	3.0	1830	US-10-750-185-38673	Sequence 38673, A
18	35.6	3.0	179487	US-10-330-773-664	Sequence 664, App
19	35.2	3.0	614	US-09-925-065A-108174	Sequence 108174,
20	35.2	3.0	616	US-09-925-065A-299192	Sequence 299192,

21	35.2	3.0	1144	US-11-096-568A-1623	Sequence 1623, Ap
22	35.2	3.0	3769	US-10-793-626-4106	Sequence 4106, Ap
23	35.2	3.0	4249	US-10-793-626-4133	Sequence 4133, Ap
24	35	2.9	201	US-10-995-561-22763	Sequence 22763, A
25	35	2.9	1085	US-10-775-169-78	Sequence 78, Appl
26	35	2.9	1233	US-11-098-686-9315	Sequence 9315, Ap
27	35	2.9	3581	US-10-750-185-52094	Sequence 52094, A
28	35	2.9	3581	US-10-750-185-52094	Sequence 52094, A
29	35	2.9	4356	US-10-932-182A-2957	Sequence 2957, Ap
30	35	2.9	4356	US-10-932-182A-2957	Sequence 2957, Ap
31	35	2.9	10144	US-10-240-708-93	Sequence 93, Appl
32	35	2.9	317876	US-10-995-561-13279	Sequence 13279, A
33	35	2.9	1457639	US-11-098-686-8739	Sequence 8739, Ap
34	34.8	2.9	550	US-09-925-065A-256291	Sequence 256291,
35	34.6	2.9	464	US-09-925-065A-383958	Sequence 383958,
36	34.6	2.9	567	US-09-925-065A-655069	Sequence 655069,
37	34.6	2.9	1400	US-11-136-527-4230	Sequence 4230, Ap
38	34.6	2.9	2273	US-11-136-527-134	Sequence 134, App
39	34.6	2.9	194186	US-10-330-773-395	Sequence 395, App
40	34.6	2.9	225587	US-10-330-773-374	Sequence 374, App
41	34.4	2.9	559	US-09-925-065A-257742	Sequence 257742,
42	34.4	2.9	623	US-09-925-065A-279572	Sequence 279572,
43	34.4	2.9	1496	US-10-750-185-63067	Sequence 63067, A
44	34.4	2.9	1496	US-10-750-185-63067	Sequence 63067, A
45	34.2	2.9	201	US-10-995-561-51601	Sequence 51601, A

ALIGNMENTS

RESULT 1
US-09-925-065A-175209/c
Sequence 175209, Application US/09925065A
GENERAL INFORMATION:
APPLICANT: Mang, David G.
TITLE OF INVENTION: Identification and Mapping of Single
FILE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
FILE REFERENCE: 108827.135
CURRENT APPLICATION NUMBER: US/09/925,065A
PRIOR FILING DATE: 2001-08-08
PRIOR APPLICATION NUMBER: US 60/243,096
PRIOR FILING DATE: 2000-10-24
PRIOR APPLICATION NUMBER: US 60/252,147
PRIOR FILING DATE: 2000-11-20
PRIOR APPLICATION NUMBER: US 60/250,092
PRIOR FILING DATE: 2000-11-30
PRIOR APPLICATION NUMBER: US 60/261,766
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: US 60/289,846
PRIOR FILING DATE: 2001-05-09
NUMBER OF SEQ ID NOS: 957086
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 175209
LENGTH: 542
TYPE: DNA
ORGANISM: Homo sapiens
US-09-925-065A-175209

Query Match 3.3%; Score 39.6; DB 6; Length 542;
Best Local Similarity 56.0%; Pred. No. 1.1;
Matches 75; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

QY 1055 TTGAATATTGAAACCTATCAATGTTTCTAGCTCAAGCTAGCTATGTCATGTC 1114
DB TTTATGATTTGAAATATTCATTTTAAAGATATGATTAAGTCTTCGCTAGGCG 325
QY 1115 TAAAGTCATGTCATTTTATTAATCACTGATCAAAACATTTGTTATGATTTTACCCAA 1174
DB 324 CAGTTCAGCTTATTCGCGAAACACATGAAAGCTTTTATTTTATTTTATTTT 265
QY 1175 AAAAAAAAAAAAAA 1188
|||||

Db 264 AAAAAAAAAAAAA 251

RESULT 2

US-09-925-065A-683561/C

Sequence 683561, Application US/09925065A

Publication No. US20040181048A1

GENERAL INFORMATION:

APPLICANT: Wang, David G.

TITLE OF INVENTION: Identification and Mapping of Single

FILE REFERENCE: 108827.135

CURRENT APPLICATION NUMBER: US/09/925,065A

PRIOR FILING DATE: 2001-08-08

PRIOR APPLICATION NUMBER: US 60/243,096

PRIOR FILING DATE: 2000-10-24

PRIOR APPLICATION NUMBER: US 60/252,147

PRIOR FILING DATE: 2000-11-20

PRIOR APPLICATION NUMBER: US 60/250,092

PRIOR FILING DATE: 2000-11-30

PRIOR APPLICATION NUMBER: US 60/261,766

PRIOR FILING DATE: 2001-01-16

PRIOR APPLICATION NUMBER: US 60/289,846

PRIOR FILING DATE: 2001-05-09

NUMBER OF SEQ ID NOS: 957086

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 683561

LENGTH: 958

TYPE: DNA

ORGANISM: Homo sapiens

US-09-925-065A-683561

Query Match 3.2%; Score 38.6; DB 6; Length 958;

Best Local Similarity 47.7%; Pred. No. 2.6; Indels 0; Gaps 0;

Matches 113; Conservative 0; Mismatches 124; Indels 0; Gaps 0;

Qy 854 AAAAGTCTGAGATGAGAGCTTTGTAGGTGGAATGTGGCTTCTTCAAGTATAGT 913

Db 938 AAAAGATATGCGGAGATGTCTATGAAGATGGATGTAGAGCTATCAAGTATAGT 879

Qy 914 TGTGGGGTGAAGAAAGCGAGATTCGATTTGTGAAGATTCCTCCGTCGCGCA 973

Db 878 GCTTGTTGGGAATTTGTTGACTTCAATGATTTAATAGATCTTATCTCTGTAG 819

Qy 974 AGCTGATCTTCACTGATGTTGAAGAGCTTCGTATCTGCTAGATTTGTGTA 1033

Db 818 TTATTCATATTCAGAGCCCATGTGGATCTTTATCTTTTGGCTGTGATTGTA 759

Qy 1034 TGTACTATTATCAGAGGGCTTTGAAATTTGAAAAACCTATCAATGTTTCTAGCT 1090

Db 758 TGTGTGTTTGTAGAGGAAGAGTGTCTCATCTGAATGTCAAGAAAAACATTTGTGTT 702

RESULT 3

US-09-925-065A-102394

Sequence 102394, Application US/09925065A

Publication No. US20040181048A1

GENERAL INFORMATION:

APPLICANT: Wang, David G.

TITLE OF INVENTION: Identification and Mapping of Single

FILE REFERENCE: 108827.135

CURRENT APPLICATION NUMBER: US/09/925,065A

PRIOR FILING DATE: 2001-08-08

PRIOR APPLICATION NUMBER: US 60/243,096

PRIOR FILING DATE: 2000-10-24

PRIOR APPLICATION NUMBER: US 60/252,147

PRIOR FILING DATE: 2000-11-20

PRIOR APPLICATION NUMBER: US 60/250,092

PRIOR FILING DATE: 2000-11-30

PRIOR APPLICATION NUMBER: US 60/261,766

PRIOR FILING DATE: 2001-01-16

PRIOR APPLICATION NUMBER: US 60/289,846

PRIOR FILING DATE: 2001-05-09

NUMBER OF SEQ ID NOS: 957086

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 102394

LENGTH: 534

TYPE: DNA

ORGANISM: Homo sapiens

US-09-925-065A-102394

Query Match 3.2%; Score 37.6; DB 6; Length 534;

Best Local Similarity 53.4%; Pred. No. 3.7; Indels 0; Gaps 0;

Matches 79; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

Qy 159 TTCAAGCAGCACCGCAAAACGGGTGAATCACTGTGCAACGTCGCGGAATTTCC 218

Db 269 TCCCAAGTGTGGAGATTACAGGTGTAGCCACCATACCAAGTATTAAGCATTTAAAC 328

Qy 219 CGTGTATTATTTCTCCAGCTTTTATCTTCCAACTACTCTGACGTTCTTAA 278

Db 329 ACTGTGTATTGTATACATGCGCTTTATTTAAATGTGTAAATGCTAAAGCCTTTTCT 388

Qy 279 CCACATCGCTTCCGATGCTTACATTTCT 306

Db 389 AAAACTAATCCCAAAAAGTACATTTCT 416

RESULT 4

US-09-925-065A-849244

Sequence 849244, Application US/09925065A

Publication No. US20040181048A1

GENERAL INFORMATION:

APPLICANT: Wang, David G.

TITLE OF INVENTION: Identification and Mapping of Single

FILE REFERENCE: 108827.135

CURRENT APPLICATION NUMBER: US/09/925,065A

PRIOR FILING DATE: 2001-08-08

PRIOR APPLICATION NUMBER: US 60/243,096

PRIOR FILING DATE: 2000-10-24

PRIOR APPLICATION NUMBER: US 60/252,147

PRIOR FILING DATE: 2000-11-20

PRIOR APPLICATION NUMBER: US 60/250,092

PRIOR FILING DATE: 2000-11-30

PRIOR APPLICATION NUMBER: US 60/261,766

PRIOR FILING DATE: 2001-01-16

PRIOR APPLICATION NUMBER: US 60/289,846

PRIOR FILING DATE: 2001-05-09

NUMBER OF SEQ ID NOS: 957086

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 849244

LENGTH: 543

TYPE: DNA

ORGANISM: Homo sapiens

US-09-925-065A-849244

Query Match 3.2%; Score 37.6; DB 6; Length 543;

Best Local Similarity 46.5%; Pred. No. 3.8; Indels 0; Gaps 0;

Matches 121; Conservative 0; Mismatches 139; Indels 0; Gaps 0;

Qy 828 TTGTATGTGTAAGATGGGCAAGAAAAAGCTGAGATGAGAGCTTTAGGTGAAT 887

Db 28 TTATATATTTGACAAACTCCCAAAATTTGAATCTTAAATTAAGTATTTGACCTAGAT 87

Qy 888 TGTGTGCGTTTCTCAAGATATGTTGTGTTGGGTGAAAAAGCGAGATTGATTTGT 947

Db 88 TGAATTTGGGATTTTCCAGTACGCGCCGTGAGAGCTCAAGAGATTAATCCCTCATTAAT 147

Qy 948 GAAGATCTTCCGTTTCTCCGCAAGCTTGAATCTTCACTGAGTTGGAAGAGCTTC 1007

Db 148 GCAAGATGTTAAAGTTATCAAGCTTGTGATGTGTGAATTTGAGGAAGATTTGTC 207

Qy 1008 TGTATCTTGTGTAGATTTGTGTATGTAATTAATCAAGAGGGCTTGAATTTTGA 1067

Db 208 AATACAGAGTGAATTAATCTTTAAATATATATTTATGATATTTTCAGAAATGCT 267
QY 1068 AACCTATCATGTTTCTTA 1087
Db 268 ATGAATTCATGAGAAATCTA 287

RESULT 5

US-09-925-065A-902549/c
; Sequence 902549, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925.065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 902549
; LENGTH: 605
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-902549

Query Match 3.1%; Score 37.4; DB 6; Length 605;
Best Local Similarity 51.5%; Pred. No. 4.5;
Matches 86; Conservative 0; Mismatches 81; Indels 0; Gaps 0;
QY 1015 TTGCTAGATTGTTGTTAGTACTATATATCAGAGGCTCTGAAATTTGAAAAACCTA 1074
Db 339 TTCTTAACCTGTTTAAATTTCTGCATTAAGCCANAGTGTTCAGAGAAATGTCGT 280
QY 1075 TCAATGTTTCTAGCTCCAGCTAGCTAATGTTTCATGTCCTAAGTGCATGTTATTTT 1134
Db 279 TAAATTCATGATATTTTACAAATTTCCAGTGTCTTGTATTTGATTTCTTGTATTTTCT 220
QY 1135 ATTAACCTGATCAAAACATTTGTTATGTTTACCCCAAAAAAAA 1181
Db 219 ATTGTGCTCGAATAATATTTGATATGTTTATTTTAAAAAAA 173

RESULT 6

US-09-925-065A-358029
; Sequence 358029, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925.065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846

; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 358029
; LENGTH: 578
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-358029

Query Match 3.1%; Score 37.2; DB 6; Length 578;
Best Local Similarity 50.0%; Pred. No. 5;
Matches 93; Conservative 0; Mismatches 93; Indels 0; Gaps 0;

QY 957 TTCGTTCTCCGCGCAAGCTGATCTTCACCTGATGGAGAACTTGTGATCTT 1016
Db 302 TTTTCTTTTGGTCTCATCTTATTTCTTATTCATTTTGGCAATCTTCTTTT 361
QY 1017 CGTCTGATTTGTTATGATATATATCAGAGGCTCTGAAATTTGAAAAACCTATC 1076
Db 362 TTTTCTTTTAAAGTGAAGATCAGTCAATGAAATTTGAACCTTTCTTGAATATCC 421
QY 1077 AATGTTTCTAGCTCCAGCTAGCTATGTTTCATGTCCTAAGTGCATGTTATTTAT 1136
Db 422 ATTTTCTTAACTTTTGTGATATCCAAATTTTGTATTTTGCATTTTCAATTTGAT 481
QY 1137 TAAACT 1142
Db 482 TCAACT 487

RESULT 7

US-09-925-065A-106455
; Sequence 106455, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925.065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 106455
; LENGTH: 590
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-106455

Query Match 3.1%; Score 37; DB 6; Length 590;
Best Local Similarity 55.8%; Pred. No. 5.7;
Matches 92; Conservative 0; Mismatches 70; Indels 3; Gaps 1;

QY 1005 TTCTGATCTTCGTTAGATTTGTTATGATATTAATCAGAGGCTCTGAAATATTT 1064
Db 305 TTCTTTAACTTGTGTGATATTTTGTCTCTTAATTTCTTAATGAGTGTGATTAATC 364
QY 1065 GAAAAACCTATCA--ATGTTTCTAGCTCCAGCTAGCTATGTTCAATGCTTAAGTG 1121
Db 365 TAAATTAATTAATACAGTATCTCCCTATTCCTAGGAGATTAATGTTCAAGAACCCCTAAG 424
QY 1122 CATGTATTTTATTAACCTGATCAAAACATTTGTTATGTTT 1166

Db 425 CATGTGTGGGGTAATTAGAGTGAACATGTGGTAATTAGAGTGT 469

RESULT 8
US-09-925-065A-417521/c
; Sequence 417521, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:

APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR FILING DATE: 2000-10-24
; PRIOR FILING DATE: 2000-11-20
; PRIOR FILING DATE: 2000-11-20
; PRIOR FILING DATE: 2000-11-30
; PRIOR FILING DATE: 2000-11-30
; PRIOR FILING DATE: 2001-01-16
; PRIOR FILING DATE: 2001-05-09
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 417521
; LENGTH: 594
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-417521

Query Match 3.1%; Score 37; DB 6; Length 594;
Best Local Similarity 55.8%; Pred. No. 5.7;
Matches 92; Conservative 0; Mismatches 70; Indels 3; Gaps 1;
Qy 1005 TTCTGTAATCTTCGCTAGATTGTTGTTATGTAATATCAAGGGCTTGAATATTT 1064
Db 482 TCCTTTAACTTTGTGTGATTTTGTGTCTCCATTTTCTTAAGGTGTGAATATC 423
Qy 1065 GAAAAACCTATCA--ATGTTTCTAGCTCCAGCTAGCTATTGTCATGTCTAAGTTG 1121
Db 422 TAAATATATATTAACGATGTTCCCTTATCCTAGGAGATATGTTCCAAAGCCCTAGTG 363
Qy 1122 CATGTGATTTTATTAACCTGATCAAAACATTTGTTATAGTTT 1166
Db 362 CATGTGTGGGGTAATTAGAGTGAACATGTGGTAATTAGAGTGT 318

RESULT 9
US-09-925-065A-106454
; Sequence 106454, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR FILING DATE: 2000-10-24
; PRIOR FILING DATE: 2000-11-20
; PRIOR FILING DATE: 2000-11-20
; PRIOR FILING DATE: 2000-11-30
; PRIOR FILING DATE: 2001-01-16
; PRIOR FILING DATE: 2001-05-09
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 106454

LENGTH: 599
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-106454

Query Match 3.1%; Score 37; DB 6; Length 599;
Best Local Similarity 55.8%; Pred. No. 5.8;
Matches 92; Conservative 0; Mismatches 70; Indels 3; Gaps 1;
Qy 1005 TTCTGTAATCTTCGCTAGATTGTTGTTATGTAATATCAAGGGCTTGAATATTT 1064
Db 241 TCCTTTAACTTTGTGTGATTTTGTGTCTCCATTTTCTTAAGGTGTGAATATC 300
Qy 1065 GAAAAACCTATCA--ATGTTTCTAGCTCCAGCTAGCTATTGTCATGTCTAAGTTG 1121
Db 301 TAAATATATATTAACGATGTTCCCTTATCCTAGGAGATATGTTCCAAAGCCCTAGTG 360
Qy 1122 CATGTGATTTTATTAACCTGATCAAAACATTTGTTATAGTTT 1166
Db 361 CATGTGTGGGGTAATTAGAGTGAACATGTGGTAATTAGAGTGT 405

RESULT 10
US-09-925-065A-283413/c
; Sequence 283413, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR FILING DATE: 2000-10-24
; PRIOR FILING DATE: 2000-11-20
; PRIOR FILING DATE: 2000-11-20
; PRIOR FILING DATE: 2000-11-30
; PRIOR FILING DATE: 2000-11-30
; PRIOR FILING DATE: 2001-01-16
; PRIOR FILING DATE: 2001-05-09
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 283413
; LENGTH: 893
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-283413

Query Match 3.1%; Score 37; DB 6; Length 893;
Best Local Similarity 50.9%; Pred. No. 6.9;
Matches 88; Conservative 0; Mismatches 85; Indels 0; Gaps 0;
Qy 995 TGAAGAAGCTTCGATATCTGCTAGATTGTTGTTATGTAATATCAAGGGCTC 1054
Db 323 TGCAATCAAAGTTAACTTTGTTAATATGAATATGATATGATATTTTACATATTT 264
Qy 1055 TTGAATATTTGAAAAACCTATCAATGTTTCTAGCTCAAGCTAGCTATTGTCATGTC 1114
Db 263 TTGATTTTMTAAAGCAATAAAGTCTTTATTTTCACTTGATTTGATTCAGATC 204
Qy 1115 TAACTTCATGATGATTTTATTAACCTGATCAAAACATTTGTTATAGTTT 1167
Db 203 TTCTGTGTTTATTTTTCCTCTGTGCTAGCTAAAGTATTTATATATTT 151

RESULT 11
US-10-240-708-29
; Sequence 29, Application US/10240708
; Publication No. US20050282157A1
; GENERAL INFORMATION:

```
APPLICANT: OLEK, Alexander
APPLICANT: PIPERBROCK, Christian
APPLICANT: BERLIN, Kurt
TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
FILE REFERENCE: 5013.1012
CURRENT APPLICATION NUMBER: US/10/240,708
PRIORITY FILING DATE: 2002-10-03
PRIORITY APPLICATION NUMBER: PCT/EP01/03971
PRIORITY FILING DATE: 2001-04-06
PRIORITY APPLICATION NUMBER: DE 10019058.8
PRIORITY FILING DATE: 2000-04-06
PRIORITY APPLICATION NUMBER: DE 10019173.8
PRIORITY FILING DATE: 2000-04-07
PRIORITY APPLICATION NUMBER: DE 10032529.7
PRIORITY FILING DATE: 2000-06-30
PRIORITY APPLICATION NUMBER: DE 10043826.1
PRIORITY FILING DATE: 2000-09-01
NUMBER OF SEQ ID NOS: 98
SEQ ID NO 29
LENGTH: 5666
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-240-708-29
```

```
Query Match 3.1%; Score 37; DB 8; Length 5666;
Best Local Similarity 47.3%; Pred. No. 16;
Matches 112; Conservative 0; Mismatches 125; Indels 0; Gaps 0;
```

```
QY 910 AGTTGCGGGGTGAAAAAGCGAGATTGATTGTTGAAGATCCCTCGTTCCG 969
DB 115 AATTAATTAAGTTTGAAGGTAAAGTTGTTAGTGTATTTTATTTATTTTAAAT 174
QY 970 GCCAAGCTGATCCCTGACCTGAGTGGAGAAGAGCTGCTGATCTGCTAGATTGT 1029
DB 175 GGTTCGCAATGTTGTTGTTGTTGTTTATTTTATTTATTTGATTTGTTGTTT 234
QY 1030 GTTATGTAATTAATCAGAGGGGCTTGAATATTTGAAAACTATCAGTCTTACG 1089
DB 235 TTTTATTTATTTATTTATTTAGTTTGTGATTTGTTGATTTATTTATTTTAA 294
QY 1090 TCCAGCTAGTATTTGTTGATGCTCTAGTTCAGTTCAGTGTATTTTAACTGAT 1146
DB 295 ATAAATTTGTTGTTATTTGATTTTATTTATTTTATTTTAAATTTATTTGAT 351
```

```
RESULT 12
US-11-091-018-1
Sequence 1, Application US/11091018
Publication No. US2005028755A1
GENERAL INFORMATION:
APPLICANT: Greteardottir, Solveig
APPLICANT: Thorleifsson, Gudmar
APPLICANT: Gulcher, Jeffrey R.
TITLE OF INVENTION: SUSCEPTIBILITY GENE FOR HUMAN STROKE;
FILE REFERENCE: 2345.2010-016
CURRENT APPLICATION NUMBER: US/11/091,018
CURRENT FILING DATE: 2005-03-25
PRIORITY APPLICATION NUMBER: PCT/US03/29906
PRIORITY FILING DATE: 2003-09-25
PRIORITY APPLICATION NUMBER: 10/255,120
PRIORITY FILING DATE: 2002-09-25
PRIORITY APPLICATION NUMBER: 10/419,723
PRIORITY FILING DATE: 2003-04-18
PRIORITY APPLICATION NUMBER: 10/650,120
PRIORITY FILING DATE: 2003-08-27
PRIORITY APPLICATION NUMBER: 10/067,514
PRIORITY FILING DATE: 2002-02-04
PRIORITY APPLICATION NUMBER: 09/811,352
PRIORITY FILING DATE: 2001-03-19
```

```
NUMBER OF SEQ ID NOS: 102
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 1691140
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(1691140)
OTHER INFORMATION: n=A,T,C or G
US-11-091-018-1
```

```
Query Match 3.1%; Score 36.4; DB 12; Length 1691140;
Best Local Similarity 56.8%; Pred. No. 2.5e+02;
Matches 67; Conservative 0; Mismatches 51; Indels 0; Gaps 0;
```

```
QY 1053 TCTTGATATTGTAAGAAACCTATCAATGTTTCTAGCTCCAGCTGCTATTGTCATGT 1112
DB 103255 TTTTGATTTTGGGCACTTTAAGTTTATTTATTTATGTCATCAGGCCAATTTATTA 103314
QY 1113 CCTAAGTTGATGTTATTTTATTTAATCTGATCAAAACATTTGTTATAGTTTACC 1170
DB 103315 CATATGTAGATCTGAGATGTTGTTTACTGCTCAGCAAAAGATTTGTTAGTTTGC 103372
```

RESULT 13

```
US-09-925-065A-881350
Sequence 881350, Application US/09925065A
Publication No. US20040181048A1
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single
FILE REFERENCE: 108827.135
CURRENT APPLICATION NUMBER: US/09/925,065A
PRIORITY FILING DATE: 2001-08-08
PRIORITY APPLICATION NUMBER: US 60/243,096
PRIORITY FILING DATE: 2000-10-24
PRIORITY APPLICATION NUMBER: US 60/252,147
PRIORITY FILING DATE: 2000-11-20
PRIORITY APPLICATION NUMBER: US 60/250,092
PRIORITY FILING DATE: 2000-11-30
PRIORITY APPLICATION NUMBER: US 60/261,766
PRIORITY FILING DATE: 2001-01-16
PRIORITY APPLICATION NUMBER: US 60/289,846
PRIORITY FILING DATE: 2001-05-09
NUMBER OF SEQ ID NOS: 957086
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 881350
LENGTH: 617
TYPE: DNA
ORGANISM: Homo sapiens
US-09-925-065A-881350
```

```
Query Match 3.0%; Score 36; DB 6; Length 617;
Best Local Similarity 50.6%; Pred. No. 11;
Matches 87; Conservative 0; Mismatches 85; Indels 0; Gaps 0;
```

```
QY 1006 TCTGATCTTCTGATAGATTGTTGTTATGTTATTTATTCAGAGGGGCTTGAATTTG 1065
DB 184 TTTGGTTTCTTCTGCTGATTTTGTGAGTCTTGTAGATTTGATATTTAGTCTTTG 243
QY 1066 AAAACCTATCAATGTTTCTAGCTCCAGCTGATTTGTTGATGTTCTTACGTTGATG 1125
DB 244 TTAGACATACAGATGAGGATTTCTCCACTGTGTGTTGTCTGTTACTTTGCTG 303
QY 1126 TGTATTTTATTTAATCTGATCAAAACATTTGTTATAGTTTACCCCAAAA 1177
DB 304 ACTATTTCTTTGCTGTGATGAAGCTTTTGGTTAGTTAATCTCCATATA 355
```

```
RESULT 14
US-10-750-185-46341/C
```

```

: Sequence 46341, Application US/10750185
: Publication No. US2005026060341
: GENERAL INFORMATION:
: APPLICANT: MMT GENOMICS, INC.
: APPLICANT: DENISE, Sue K.
: APPLICANT: KERR, Richard
: APPLICANT: ROSENFELD, David
: APPLICANT: HOLM, Tom
: APPLICANT: BATES, Stephen
: APPLICANT: FANTIN, Dennis
: TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
: FILE REFERENCE: MM1100-2
: CURRENT APPLICATION NUMBER: US/10/750,185
: CURRENT FILING DATE: 2003-12-31
: PRIOR APPLICATION NUMBER: US 60/437,482
: PRIOR FILING DATE: 2002-12-31
: NUMBER OF SEQ ID NOS: 64922
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 46341
: LENGTH: 1462
: TYPE: DNA
: ORGANISM: Bovine 19866880447381
: US-10-750-185-46341

```

Query Match	3.0%	Score 36;	DB 8;	Length 1462;
Best Local Similarity	48.5%	Pred. No. 16;		
Matches 99; Conservative	0;	Mismatches 105;	Indels 0;	Gaps 0;

QY	977	TTGATCCCTCACCCTGAGTTGGAGAAAGCTTCGGATCTTCGCTCAGATTTGTGTATGT	1036
QY	977	TTGATCCCTCACCCTGAGTTGGAGAAAGCTTCGGATCTTCGCTCAGATTTGTGTATGT	1036
Db	937	TTTTTCCCTTACTTCACGTCTAAGTACCGACTTATTACTTTTAAATTTAGAGAAAG	878
QY	1037	ACTATTATCAGAGGGGCTTGAAATATTTGAAAAACCTATCATGTTTTCTAGCTCAAGC	10986
Db	877	AATTTGAATTTTAGAAGGGTGACATGTTTATGACTTCTATTGTTTCTTACCAAGC	818
QY	1097	TAGCTATTGTCATGTCCTAAAGTGCATGTCGTAATTTTATTAACTCGATCAAAACATTT	11586
Db	817	TGGCATTTTTTAAAGTCCCTAAATATCATCTCGACCTTAAGAACTAATCGAAGACCTT	758
QY	1157	GTTATAGTTTACCCCAAAAAA	1180
Db	757	CTTTAAGCATTCTCAAAATTTAA	734

```

RESULT 15
US-10-750-623-46341/C
Sequence 46341. Application US/10750623
Publication NO. US20050287531A1
GENERAL INFORMATION:
APPLICANT: MMI GENOMICS, INC.
APPLICANT: DENISE, Sue K.
APPLICANT: KERR, Richard
APPLICANT: ROSENFELD, David
APPLICANT: HOLM, Tom
APPLICANT: BATES, Stephen
APPLICANT: FANTIN, Dennis
TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
FILE REFERENCE: MM1100-1
CURRENT APPLICATION NUMBER: US/10/750,623
CURRENT FILING DATE: 2003-12-31
PRIOR APPLICATION NUMBER: US 60/437,482
PRIOR FILING DATE: 2002-12-31
NUMBER OF SEQ ID NOS: 64922
SOFTWARE: PatentIn version 3.1
SEQ ID NO 46341
LENGTH: 1462
TYPE: DNA
ORGANISM: Bovine 19866880447381
US-10-750-623-46341

```

Query Match	3.0%;	Score 36;	DB 8;	Length 1462;
Best Local Similarity	48.5%;	Pred. No. 16;		

	Matches	99;	Conservative	0;	Mismatches	105;	Indels	0;	Gaps	0;
Qy	977	TTGATCCTTCA	CCGAGTTGG	AAAGAA	CACTTCGTGTATCTTCGTAAATTTGCTGTATGT	1036				
Db	937	TTTTTCCCTT	TACTTCAC	AGCTAGTA	TAACCCGACTTTATTA	CTTTTAAATTTAGAGAAAG	878			
Qy	1037	ACTATTTAC	AGAGGGCTCT	TGAATATTTG	AAAAAACC	TATCAATGTTTTCT	AGCTCAAGC	1096		
Db	877	AAATTTGAT	TTTTTAC	CAAGGGTG	ACATGTTATATG	ACTTGCTATTTGTTTTCT	ACCAAGC	818		
Qy	1097	TAGCTATTTG	TTCANCTCT	TAAGTGC	ATGTATTTTATTA	AACTGCATCAAA	CAATT	1156		
Db	817	TGGCAATTTT	TAAAGTCTT	AAAAATAT	CACTCGGAC	TTTAAAGAACTA	ATCGAAGCACTT	758		
Qy	1157	GTTATAGCTT	TATACCCCA	AAAAAAA	1180					
Db	757	CTTTAGCACT	CTTCAAAAT	TTTAA	734					

```
Search completed: March 20, 2006, 14:23:26
Job time : 477.098 secs
```

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 20, 2006, 13:57:59 ; Search time 236.782 Seconds

(without alignments)
8918.522 Million cell updates/sec

Title: US-10-634-548-18

Perfect score: 1188
Sequence: 1 gatacataaattctcaacac.....cccccaaaaaaaaaaaaaa 1188

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA: *
1: /cgn2_6/ptodata/1/ina/1 COMB.seq: *
2: /cgn2_6/ptodata/1/ina/5 COMB.seq: *
3: /cgn2_6/ptodata/1/ina/6A COMB.seq: *
4: /cgn2_6/ptodata/1/ina/6B COMB.seq: *
5: /cgn2_6/ptodata/1/ina/H COMB.seq: *
6: /cgn2_6/ptodata/1/ina/PTUS COMB.seq: *
7: /cgn2_6/ptodata/1/ina/PP COMB.seq: *
8: /cgn2_6/ptodata/1/ina/RE COMB.seq: *
9: /cgn2_6/ptodata/1/ina/backfile1.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	41.6	3.5	7218	2	US-08-232-463-14
2	37.4	3.1	601	3	US-09-949-016-41156
3	37.4	3.1	601	3	US-09-949-016-41157
4	37.4	3.1	161900	3	US-09-949-016-12685
5	37.4	3.1	161914	3	US-09-949-016-12906
6	36.6	3.1	100836	3	US-09-949-016-12871
7	36.6	3.1	100837	3	US-09-949-016-17063
8	36.4	3.1	317366	3	US-09-949-016-16001
9	35.6	3.0	404	3	US-09-640-211A-2094
10	35.4	3.0	57638	3	US-09-949-016-17000
11	35.2	3.0	3769	3	US-09-710-379-4106
12	35.2	3.0	4249	3	US-09-710-379-4133
13	35.2	2.9	124110	3	US-09-949-016-13353
14	34.6	2.9	862	2	US-08-222-719-17
15	34.6	2.9	862	2	US-08-470-925-17
16	34.6	2.9	862	2	US-08-471-613-17
17	34.6	2.9	862	6	PCT-US93-10443-17
18	34.6	2.9	5520	3	US-10-001-887-43
19	34.6	2.9	11168	3	US-09-434-840-5
20	34.6	2.9	96690	3	US-09-662-2548-24
21	34.6	2.9	96690	3	US-09-949-016-17103
22	34.4	2.9	832	3	US-09-621-976-2813
23	34.4	2.9	855	3	US-09-060-726A-1
24	34.4	2.9	855	3	US-09-060-726A-1

25	34.4	2.9	856	3	US-09-060-726A-3	Sequence 3, Appl1
26	34.4	2.9	856	3	US-09-845-849A-1	Sequence 1, Appl1
27	34.4	2.9	856	3	US-09-845-849A-3	Sequence 3, Appl1
28	34.4	2.9	4032	2	US-08-107-748-3	Sequence 3, Appl1
29	34.4	2.9	4032	2	US-08-245-809-4	Sequence 4, Appl1
30	34.4	2.9	4032	6	PCT-US92-01385-3	Sequence 3, Appl1
31	34.4	2.9	164061	3	US-09-949-016-17422	Sequence 17422, A
32	34.4	2.9	264665	3	US-09-949-016-13747	Sequence 13747, A
33	34.2	2.9	145812	3	US-09-949-016-15698	Sequence 15698, A
34	34.2	2.9	640681	3	US-09-790-988-1	Sequence 1, Appl1
35	34	2.9	285	3	US-09-270-767-25614	Sequence 25614, A
36	34	2.9	1091	3	US-09-270-767-10251	Sequence 10251, A
37	34	2.9	102520	3	US-09-949-016-17367	Sequence 17367, A
38	34	2.9	102526	3	US-09-949-016-12448	Sequence 12448, A
39	33.8	2.8	420	3	US-09-513-999C-32555	Sequence 32555, A
40	33.8	2.8	523	3	US-09-270-767-4870	Sequence 4870, Ap
41	33.8	2.8	523	3	US-09-270-767-20152	Sequence 20152, A
42	33.8	2.8	33353	3	US-09-949-016-17285	Sequence 17285, A
43	33.8	2.8	40000	3	US-09-780-049-18	Sequence 18, Appl1
44	33.8	2.8	89220	3	US-09-949-016-12655	Sequence 12655, A
45	33.8	2.8	89224	3	US-09-949-016-15572	Sequence 15572, A

ALIGNMENTS

RESULT 1
US-08-232-463-14/c
Sequence 14, Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F
APPLICANT: SCHREIFLINER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-9300
TELEFAX: (703) 683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: PTZ9pt-Fls

US-08-232-463-14

Query Match	3.5%;	Score 41.6;	DB 2;	Length 7218;
Best Local Similarity	2.9%;	Pred. No. 0.055;		
Matches 11;	Conservative 212;	Mismatches 161;	Indels 0;	Gaps 0;

QY	572	AATCATTAAGAACCGATCCGGATATCTTAACGTAATTAACCGAATCTTCGAGCGGACA	631
Db	1447	AAGAATTGGTACRRR	1388
QY	632	TACCGTTGCAGTGGTGGGAACCGGACTCGAACCGAATGGAACAACGTGATGCCACCAT	691
Db	1387	RRR	1328
QY	692	GCGCAACCAACGACTTAACCATGACGAGTTTACAAGAGTGTAGCGCAGAAAGCC	751
Db	1327	RRR	1268
QY	752	ATTTCGGGTCGGGATTAACGACATATGATATGTTGACAGATGATTGCGCGTTTG	811
Db	1267	RRR	1208
QY	812	TTGGATTATGCGCGGTGTATGTGTAAAGATGGCAAGAAAAGTCTAGATGAGA	871
Db	1207	RRR	1148
QY	872	GCTTTGATGTCGAATTGTGGTTCGCGTTCTCAAGTATAGTTGTGGGGTAAAAACGG	931
Db	1147	RRR	1088
QY	932	AGATTCAATTGATGTGAAGATC	955
Db	1087	RRRRRRRRRRRRRRRRRRATC	1064

```

RESULT 2
US-09-949-016-41156
; Sequence 41156 Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 41156
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-41156

```

[illegible]

Db 451 CTAGATGAATAATTGAGCTTCTGTTATTATTAAGCCTTACCCCATCTATAA 501

```

RESULT 3
US-09-949-016-41157
: Sequence 41157, Application US/09949016
: Patent No. 6812339
: GENERAL INFORMATION:
: APPLICANT: VENTER, J. Craig et al.
: TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
: TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
: FILE REFERENCE: C0001307
: CURRENT APPLICATION NUMBER: US/09/949,016
: CURRENT FILING DATE: 2000-04-14
: PRIOR APPLICATION NUMBER: 60/241,755
: PRIOR FILING DATE: 2000-10-20
: PRIOR APPLICATION NUMBER: 60/237,768
: PRIOR FILING DATE: 2000-10-03
: PRIOR APPLICATION NUMBER: 60/231,498
: PRIOR FILING DATE: 2000-09-08
: NUMBER OF SEQ ID NOS: 207012
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 41157
: LENGTH: 601
: TYPE: DNA
: ORGANISM: Human
US-09-949-016-41157

```

Query Match	3.1%	Pred. 37.4	DB 3	Length 601
Best Local Similarity	55.0%	Pred. No. 0.29		
Matches	94	Conservative	0	Mismatches 76; Indels 1; Gaps 1
QY	1011	TATCTGCTCAGATTGTGTATGTAATATATATACAGAGGGCTTGAAATATTTGAAAA	1069	
Db	349	TATTTTGTCTAACTTTATATATTTTAAACGTAAGACGTTAAATTTTGTCTAAA	408	
QY	1070	ACCTATCAATGTTTTCAGTCCAGCTAGCTATTGTTCAATGCTCAAGTTGCATGTGA	1129	
Db	409	AACATCAATTTTATTTTCTATCAAAAGATGTTTCTTATGTTAAGTTTAA	468	
QY	1130	TTTTTATTAACGTGCATCAAAAACATTTGTTATAGTTTAACCCCAAAAAA	1180	
Db	469	CTAAGTAAAAATTTGAGGCTCTGTTTATTAAGCCCTTAAACCCCATCTATAA	519	

```

1      RESULT 4
2      US-09-949-016-12685/c
3      ; Sequence 12685, Application US/09949016
4      ; Patent No. 6812339
5      ; GENERAL INFORMATION:
6      ; APPLICANT: VENTER, J. Craig et al.
7      ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
8      ; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
9      ; FILE REFERENCE: CLO01307
10     ; CURRENT APPLICATION NUMBER: US/09/949.016
11     ; CURRENT FILING DATE: 2000-04-14
12     ; PRIOR APPLICATION NUMBER: 60/241,755
13     ; PRIOR FILING DATE: 2000-10-20
14     ; PRIOR APPLICATION NUMBER: 60/237,768
15     ; PRIOR FILING DATE: 2000-10-03
16     ; PRIOR APPLICATION NUMBER: 60/231,498
17     ; PRIOR FILING DATE: 2000-09-08
18     ; NUMBER OF SEQ ID NOS: 207012
19     ; SOFTWARE: FastSeq for Windows Version 4.0
20     ; SEQ ID NO 12685
21     ; LENGTH: 161900
22     ; TYPE: DNA
23     ; ORGANISM: Human
24     ; FEATURE:
25     ; NAME/KEY: misc_feature
26     ; LOCATION: (1)..(161900)
27     ; OTHER INFORMATION: n = A,T,C or G
28     ;
29     ;
30     ;
31     ;
32     ;
33     ;
34     ;
35     ;
36     ;
37     ;
38     ;
39     ;
40     ;
41     ;
42     ;
43     ;
44     ;
45     ;
46     ;
47     ;
48     ;
49     ;
50     ;
51     ;
52     ;
53     ;
54     ;
55     ;
56     ;
57     ;
58     ;
59     ;
60     ;
61     ;
62     ;
63     ;
64     ;
65     ;
66     ;
67     ;
68     ;
69     ;
70     ;
71     ;
72     ;
73     ;
74     ;
75     ;
76     ;
77     ;
78     ;
79     ;
80     ;
81     ;
82     ;
83     ;
84     ;
85     ;
86     ;
87     ;
88     ;
89     ;
90     ;
91     ;
92     ;
93     ;
94     ;
95     ;
96     ;
97     ;
98     ;
99     ;
100    ;
101    ;
102    ;
103    ;
104    ;
105    ;
106    ;
107    ;
108    ;
109    ;
110    ;
111    ;
112    ;
113    ;
114    ;
115    ;
116    ;
117    ;
118    ;
119    ;
120    ;
121    ;
122    ;
123    ;
124    ;
125    ;
126    ;
127    ;
128    ;
129    ;
130    ;
131    ;
132    ;
133    ;
134    ;
135    ;
136    ;
137    ;
138    ;
139    ;
140    ;
141    ;
142    ;
143    ;
144    ;
145    ;
146    ;
147    ;
148    ;
149    ;
150    ;
151    ;
152    ;
153    ;
154    ;
155    ;
156    ;
157    ;
158    ;
159    ;
160    ;
161    ;
162    ;
163    ;
164    ;
165    ;
166    ;
167    ;
168    ;
169    ;
170    ;
171    ;
172    ;
173    ;
174    ;
175    ;
176    ;
177    ;
178    ;
179    ;
180    ;
181    ;
182    ;
183    ;
184    ;
185    ;
186    ;
187    ;
188    ;
189    ;
190    ;
191    ;
192    ;
193    ;
194    ;
195    ;
196    ;
197    ;
198    ;
199    ;
200    ;
201    ;
202    ;
203    ;
204    ;
205    ;
206    ;
207    ;
208    ;
209    ;
210    ;
211    ;
212    ;
213    ;
214    ;
215    ;
216    ;
217    ;
218    ;
219    ;
220    ;
221    ;
222    ;
223    ;
224    ;
225    ;
226    ;
227    ;
228    ;
229    ;
230    ;
231    ;
232    ;
233    ;
234    ;
235    ;
236    ;
237    ;
238    ;
239    ;
240    ;
241    ;
242    ;
243    ;
244    ;
245    ;
246    ;
247    ;
248    ;
249    ;
250    ;
251    ;
252    ;
253    ;
254    ;
255    ;
256    ;
257    ;
258    ;
259    ;
260    ;
261    ;
262    ;
263    ;
264    ;
265    ;
266    ;
267    ;
268    ;
269    ;
270    ;
271    ;
272    ;
273    ;
274    ;
275    ;
276    ;
277    ;
278    ;
279    ;
280    ;
281    ;
282    ;
283    ;
284    ;
285    ;
286    ;
287    ;
288    ;
289    ;
290    ;
291    ;
292    ;
293    ;
294    ;
295    ;
296    ;
297    ;
298    ;
299    ;
300    ;
301    ;
302    ;
303    ;
304    ;
305    ;
306    ;
307    ;
308    ;
309    ;
310    ;
311    ;
312    ;
313    ;
314    ;
315    ;
316    ;
317    ;
318    ;
319    ;
320    ;
321    ;
322    ;
323    ;
324    ;
325    ;
326    ;
327    ;
328    ;
329    ;
330    ;
331    ;
332    ;
333    ;
334    ;
335    ;
336    ;
337    ;
338    ;
339    ;
340    ;
341    ;
342    ;
343    ;
344    ;
345    ;
346    ;
347    ;
348    ;
349    ;
350    ;
351    ;
352    ;
353    ;
354    ;
355    ;
356    ;
357    ;
358    ;
359    ;
360    ;
361    ;
362    ;
363    ;
364    ;
365    ;
366    ;
367    ;
368    ;
369    ;
370    ;
371    ;
372    ;
373    ;
374    ;
375    ;
376    ;
377    ;
378    ;
379    ;
380    ;
381    ;
382    ;
383    ;
384    ;
385    ;
386    ;
387    ;
388    ;
389    ;
390    ;
391    ;
392    ;
393    ;
394    ;
395    ;
396    ;
397    ;
398    ;
399    ;
400    ;
401    ;
402    ;
403    ;
404    ;
405    ;
406    ;
407    ;
408    ;
409    ;
410    ;
411    ;
412    ;
413    ;
414    ;
415    ;
416    ;
417    ;
418    ;
419    ;
420    ;
421    ;
422    ;
423    ;
424    ;
425    ;
426    ;
427    ;
428    ;
429    ;
430    ;
431    ;
432    ;
433    ;
434    ;
435    ;
436    ;
437    ;
438    ;
439    ;
440    ;
441    ;
442    ;
443    ;
444    ;
445    ;
446    ;
447    ;
448    ;
449    ;
450    ;
451    ;
452    ;
453    ;
454    ;
455    ;
456    ;
457    ;
458    ;
459    ;
460    ;
461    ;
462    ;
463    ;
464    ;
465    ;
466    ;
467    ;
468    ;
469    ;
470    ;
471    ;
472    ;
473    ;
474    ;
475    ;
476    ;
477    ;
478    ;
479    ;
480    ;
481    ;
482    ;
483    ;
484    ;
485    ;
486    ;
487    ;
488    ;
489    ;
490    ;
491    ;
492    ;
493    ;
494    ;
495    ;
496    ;
497    ;
498    ;
499    ;
500    ;
501    ;
502    ;
503    ;
504    ;
505    ;
506    ;
507    ;
508    ;
509    ;
510    ;
511    ;
512    ;
513    ;
514    ;
515    ;
516    ;
517    ;
518    ;
519    ;
520    ;
521    ;
522    ;
523    ;
524    ;
525    ;
526    ;
527    ;
528    ;
529    ;
530    ;
531    ;
532    ;
533    ;
534    ;
535    ;
536    ;
537    ;
538    ;
539    ;
540    ;
541    ;
542    ;
543    ;
544    ;
545    ;
546    ;
547    ;
548    ;
549    ;
550    ;
551    ;
552    ;
553    ;
554    ;
555    ;
556    ;
557    ;
558    ;
559    ;
560    ;
561    ;
562    ;
563    ;
564    ;
565    ;
566    ;
567    ;
568    ;
569    ;
570    ;
571    ;
572    ;
573    ;
574    ;
575    ;
576    ;
577    ;
578    ;
579    ;
580    ;
581    ;
582    ;
583    ;
584    ;
585    ;
586    ;
587    ;
588    ;
589    ;
590    ;
591    ;
592    ;
593    ;
594    ;
595    ;
596    ;
597    ;
598    ;
599    ;
600    ;
601    ;
602    ;
603    ;
604    ;
605    ;
606    ;
607    ;
608    ;
609    ;
610    ;
611    ;
612    ;
613    ;
614    ;
615    ;
616    ;
617    ;
618    ;
619    ;
620    ;
621    ;
622    ;
623    ;
624    ;
625    ;
626    ;
627    ;
628    ;
629    ;
630    ;
631    ;
632    ;
633    ;
634    ;
635    ;
636    ;
637    ;
638    ;
639    ;
640    ;
641    ;
642    ;
643    ;
644    ;
645    ;
6
```

US-09-949-016-12685

Query Match 3.1%; Score 37.4; DB 3; Length 161900;
Best Local Similarity 55.0%; Pred. No. 7.2;
Matches 94; Conservative 0; Mismatches 76; Indels 1; Gaps 1;

QY 1011 TATTCGTCTAGATTTGTGTATGCTACTATATACAGAGGGCTTGAAAT-TTGAAA 1069
DB 81619 TATTTTGTCTAACCTTATATGATTTTAAACGTAAGAGCTTAAATTTTGTGTAAA 81560
QY 1070 ACCTATCAATGTTTCTAGCTCCAGAGCTAGCTATTTGCTCAAGTTGATGTGTA 1129
DB 81559 AACATCAATTTTATTTTCCATCAAAAGATTTGATTTTCTTATGTTAGTTTAA 81500
QY 1130 TTTTATTAACCTGATCAAAACATTTGTTATAGTTTACCCTCAAAAAA 1180
DB 81499 CTAGATCAAAATTTGAGGTCTGTTTATTTAAGCTTTACCCCATCTATA 81449

RESULT 5

US-09-949-016-12906/c
Sequence 12906, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:

APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
FILE REFERENCE: C1001307
CURRENT APPLICATION NUMBER: US/09/949, 016
PRIOR FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 12906
LENGTH: 161914
TYPE: DNA
ORGANISM: Human
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(161914)
OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12906

Query Match 3.1%; Score 37.4; DB 3; Length 161914;
Best Local Similarity 55.0%; Pred. No. 7.2;
Matches 94; Conservative 0; Mismatches 76; Indels 1; Gaps 1;

QY 1011 TATTCGTCTAGATTTGTGTATGCTACTATATACAGAGGGCTTGAAAT-TTGAAA 1069
DB 81629 TATTTTGTCTAACCTTATATGATTTTAAACGTAAGAGCTTAAATTTTGTGTAAA 81570
QY 1070 ACCTATCAATGTTTCTAGCTCCAGAGCTAGCTATTTGCTCAAGTTGATGTGTA 1129
DB 81569 AACATCAATTTTATTTTCCATCAAAAGATTTGATTTTCTTATGTTAGTTTAA 81510
QY 1130 TTTTATTAACCTGATCAAAACATTTGTTATAGTTTACCCTCAAAAAA 1180
DB 81509 CTAGATCAAAATTTGAGGTCTGTTTATTTAAGCTTTACCCCATCTATA 81459

RESULT 6

US-09-949-016-12871/c
Sequence 12871, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:

APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
OTHER INFORMATION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

FILE REFERENCE: C1001307
CURRENT APPLICATION NUMBER: US/09/949, 016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 12871
LENGTH: 100836
TYPE: DNA
ORGANISM: Human
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(100836)
OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12871

Query Match 3.1%; Score 36.6; DB 3; Length 100836;
Best Local Similarity 48.3%; Pred. No. 9.9;
Matches 102; Conservative 0; Mismatches 109; Indels 0; Gaps 0;

QY 958 TCCGTTCTCCGGCCAGAGCTTGATCCTTACCTGAGTTGAGAGAGCTTCTGTATCTTC 1017
DB 35075 TCAACTTCCCAATTTACCTTATCCAAATCTGCTTCCAAATCTTTGGACTCAGATT 35016
QY 1018 GTCTAATTTGTCTTATGCTATTTATACAGAGGGCTTGAATTTTGAATTCATCA 1077
DB 35015 CAGAGATTTTATATTTTCTGATTTTATATTTATTTTAAATTTTCAATATTAAC 34956
QY 1078 ATGTTTCTAGCTCCAGAGCTAGCTATTTGCTCAAGTTGATGTATTTTATTTAT 1137
DB 34955 TTTTATCTGACCCCTTATGCTCCTTTATATGAAATTTTGAAGAAATGATCA 34896
QY 1138 AAATCGATCAAAACATTTGTTATAGTTTAA 1168
DB 34895 ACACCTTCTCTCAACAGATGTATATATTA 34865

RESULT 7

US-09-949-016-17063/c
Sequence 17063, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:

APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
FILE REFERENCE: C1001307
CURRENT APPLICATION NUMBER: US/09/949, 016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 17063
LENGTH: 100837
TYPE: DNA
ORGANISM: Human
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(100837)
OTHER INFORMATION: n = A,T,C or G
US-09-949-016-17063

Query Match 3.1%; Score 36.6; DB 3; Length 100837;
Best Local Similarity 48.3%; Pred. No. 9.9;

Matches 102; Conservative 0; Mismatches 109; Indels 0; Gaps 0;

```
OY 958 TCCGTTTCGCCGCAAGCTTGATTCCTCACTGAGTGGAGAACCTTCGTATCTTC 1017
DB 35075 TCAACTTCCCAATTACCTTATCAATCTCTGCTCCATTCCTTGACATCAGATT 35016
OY 1018 GTCTGATTTGGTGTATGACTATATATGAGAGGGGCTTGAAATTTGAAAACTATCA 1077
DB 35015 CAGAAATTTTATATTTTGTGATTTTATATATTTCTTTATATATTTCAATATTAAC 34956
OY 1078 ATGTTTCTAGCTCCAGCTACGATGTTGCTATGCTTCAAGTGGATGATTTTATT 1137
DB 34955 TTTTATCTGACCCCTATGCTCCCTTATATGAAATTTTGAGAAATGACATATGA 34896
OY 1138 AAACGATCAAAACATTTGTTATAGTTTA 1168
DB 34895 ACATCTTCTCTCAAAAGTTATATATTA 34865
```

RESULT 8

US-09-949-016-16001
; Sequence 16001, Application US/09949016
; Patent No. 6812339

; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16001
; LENGTH: 317366
; TYPE: DNA
; ORGANISM: Human
; NAME/KEY: misc feature
; LOCATION: (1)_(317366)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-16001

Query Match 3.1%; Score 36.4; DB 3; Length 317366;
Best Local Similarity 49.5%; Pred. No. 22;
Matches 94; Conservative 0; Mismatches 96; Indels 0; Gaps 0;

```
OY 978 TGATCCTTCACCTGAGTGGAGAAAGCTTGCTGATCTTCGTAGATTGTGTATGTA 1037
DB 151655 TTATTCATTAATTTTGAATTTATTCATGATTCATTTTAAATGTGTTCTTTA 151714
OY 1038 CTATATCAGAGGGGCTTGAAATTTGAAAACTATCAAAAGTTTCTAGTCCAAGCT 1097
DB 151715 ATTAATATATTTTAAAGTATTTTACAAAACCTTACCAATTTCTTTTTTTTAA 151774
OY 1098 AGCTATTGTTCATGCTTCAAGTGTGATTTTATTAATCACTGATCAAAACATTTG 1157
DB 151775 ACTTTATTTTAGTTCAGGGGTACATGTGACGTTTGTATATAGTATACCTTATGTC 151834
OY 1158 TTATAGTTTT 1167
DB 151835 ATGGGTTTT 151844
```

RESULT 9

US-09-640-211A-2094
; Sequence 2094, Application US/09640211A
; Patent No. 6833446

; GENERAL INFORMATION:

; APPLICANT: Wood, Marion
; APPLICANT: Shenk, Michael A.
; APPLICANT: McGrath, Annette
; APPLICANT: Glenn, Matthew
; TITLE OF INVENTION: Compositions and Methods for the
; FILE REFERENCE: 11000.1021CIU
; CURRENT APPLICATION NUMBER: US/09/640,211A
; CURRENT FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 2368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2094
; LENGTH: 404
; TYPE: DNA
; ORGANISM: Pinus radiata
US-09-640-211A-2094

Query Match 3.0%; Score 35.6; DB 3; Length 404;
Best Local Similarity 51.2%; Pred. No. 0.86;
Matches 83; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

```
OY 559 GCAGAACTAACAAATACATTAGAACCGATCCGATATCTTAACGTATAACCGGAATCT 618
DB 219 GAAGAAAGTGACAGAGTACATCAAAAGACAGATGCTCCAGTGTATTAACGAGAAAGT 278
OY 619 TTGAGCTGACATACCGGTTGACGTGGGAAACCGACTCGAGCCGAAGTGGAAACAC 678
DB 279 GCGAATATGAGAAAGAGATTGAATTTCTGAATCCATTCGAAAGAAAGATGTGGCGAG 338
OY 679 GTGATGCCACCATCGACCAACCGACTTAACCATGAGAG 720
DB 339 GAGCTGGCATCATGTGATGAAAGATTGATGATGAGTGGAG 380
```

RESULT 10

US-09-640-211A-354
; Sequence 354, Application US/09640211A
; Patent No. 6833446

; GENERAL INFORMATION:
; APPLICANT: Wood, Marion
; APPLICANT: Shenk, Michael A.
; APPLICANT: McGrath, Annette
; APPLICANT: Glenn, Matthew
; TITLE OF INVENTION: Compositions and Methods for the
; FILE REFERENCE: 11000.1021CIU
; CURRENT APPLICATION NUMBER: US/09/640,211A
; CURRENT FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 2368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 354
; LENGTH: 405
; TYPE: DNA
; ORGANISM: Pinus radiata
US-09-640-211A-354

Query Match 3.0%; Score 35.6; DB 3; Length 405;
Best Local Similarity 51.2%; Pred. No. 0.86;
Matches 83; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

```
OY 559 GCAGAACTAACAAATACATTAGAACCGATCCGATATCTTAACGTATAACCGGAATCT 618
DB 219 GAAGAAAGTGACAGAGTACATCAAAAGACAGATGCTCCAGTGTATTAACGAGAAAGT 278
OY 619 TTGAGCTGACATACCGGTTGACGTGGGAAACCGACTCGAGCCGAAGTGGAAACAC 678
DB 279 GCGAATATGAGAAAGAGATTGAATTTCTGAATCCATTCGAAAGAAAGATGTGGCGAG 338
OY 679 GTGATGCCACCATCGACCAACCGACTTAACCATGAGAG 720
DB 339 GAGCTGGCATCATGTGATGAAAGATTGATGATGAGTGGAG 380
```

RESULT 11
US-09-949-016-17000/C
Sequence 17000, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIORITY APPLICATION NUMBER: 60/241,755
PRIORITY FILING DATE: 2000-10-20
PRIORITY APPLICATION NUMBER: 60/237,768
PRIORITY FILING DATE: 2000-10-03
PRIORITY APPLICATION NUMBER: 60/231,498
PRIORITY FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: ParseSeq for Windows Version 4.0
SEQ ID NO 17000
LENGTH: 57638
TYPE: DNA
ORGANISM: Human
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(57638)
OTHER INFORMATION: n = A,T,C or G
US-09-949-016-17000

Query Match
Best Local Similarity 3.0%; Score 35.4; DB 3; Length 57638;
Matches 63; Conservative 0; Mismatches 46; Indels 0; Gaps 0;
QY 1050 GGGTCTGATATATGAAAAAAGCTATCATGTTTCTAGCTCAAGCTAGTATTGTCA 1109
DB 19796 GGGATGGGACACAGTTTAAACACAAATTCATTATTTCAACATCTTATGCATA 19737
QY 1110 TGTCTTAAGTTGCATGTTGTTTATTTAACTGCATCAAAACATTTGT 1158
DB 19736 CAGCTGAATGTATGTATTTATTAACAATTAAAAATTAATTTGT 19688

RESULT 12
US-09-710-279-4106/C
Sequence 4106, Application US/09710279
Patent No. 6703492
GENERAL INFORMATION:
APPLICANT: KIMBERLY, WILLIAM JOHN
TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
FILE REFERENCE: P03480US
CURRENT APPLICATION NUMBER: US/09/710,279
CURRENT FILING DATE: 2000-11-09
PRIORITY APPLICATION NUMBER: 60/164,258
PRIORITY FILING DATE: 1999-11-09
NUMBER OF SEQ ID NOS: 4472
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 4106
LENGTH: 3769
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: synthetic
OTHER INFORMATION: nucleic acid sequence
US-09-710-279-4106

Query Match
Best Local Similarity 3.0%; Score 35.2; DB 3; Length 3769;
Matches 91; Conservative 0; Mismatches 93; Indels 0; Gaps 0;
QY 1001 AAGCTTCTGATCTTCTGATAGTTTGTATTGATCTATTTATCAGAGGGCTTGAAT 1060
DB 3181 AATGTTGATATCATCAATTCATTAATGTATGTATGTATTAACGAATGTTTAAA 3122

QY 1061 ATTGAAAAACATCAATGTTTCTAGCTCAAGCTAGCTATGTTCAATGCTTAAGTT 1120
DB 3121 AGTGAATTTGTTCTTTTATTTTAAACATTTGATTAATGAAACATGTTTAAATGTT 3062
QY 1121 GCATGCTATTTTATTAACCTGCATCAAAACATTTGTTATGTTTACCCCAAAAAA 1180
DB 3061 GAATCTTTTATGATTTTAAATGATTAATTAATTAATGATTTTAAAGCTTAGACATA 3002
QY 1181 AAAA 1184
DB 3001 CAAA 2998

RESULT 13
US-09-710-279-4133/C
Sequence 4133, Application US/09710279
Patent No. 6703492
GENERAL INFORMATION:
APPLICANT: KIMBERLY, WILLIAM JOHN
TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
FILE REFERENCE: P03480US
CURRENT APPLICATION NUMBER: US/09/710,279
CURRENT FILING DATE: 2000-11-09
PRIORITY APPLICATION NUMBER: 60/164,258
PRIORITY FILING DATE: 1999-11-09
NUMBER OF SEQ ID NOS: 4472
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 4133
LENGTH: 4249
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: synthetic
OTHER INFORMATION: nucleic acid sequence
US-09-710-279-4133

Query Match
Best Local Similarity 3.0%; Score 35.2; DB 3; Length 4249;
Matches 91; Conservative 0; Mismatches 93; Indels 0; Gaps 0;
QY 1001 AAGCTTCTGATCTTCTGATAGTTTGTATTGATCTATTTATCAGAGGGCTTGAAT 1060
DB 1763 AATGTTGATATCATCAATTCATTAATGTATGTATTAATGAAACATGTTTAAATGTT 1704
QY 1061 ATTGAAAAACATCAATGTTTCTAGCTCAAGCTAGCTATGTTCAATGCTTAAGTT 1120
DB 1703 AGTGAATTTGTTCTTTTATTTTAAACATTTGATTAATGAAACATGTTTAAATGTT 1644
QY 1121 GCATGCTATTTTATTAACCTGCATCAAAACATTTGTTATGTTTACCCCAAAAAA 1180
DB 1643 GAATCTTTTATGATTTTAAATGATTAATTAATTAATGATTTTAAAGCTTAGACATA 1584
QY 1181 AAAA 1184
DB 1583 CAAA 1580

RESULT 14
US-09-949-016-13353
Sequence 13353, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIORITY APPLICATION NUMBER: 60/241,755
PRIORITY FILING DATE: 2000-10-20
PRIORITY APPLICATION NUMBER: 60/237,768
PRIORITY FILING DATE: 2000-10-03

Result No.	Score	Query Match	length	DB	ID	Description
1	1135	100.0	1135	6	E63072	DNA encoding
2	1135	100.0	1135	15	AF134302	Arabidop
3	998.2	87.9	1025	15	BT002898	Arabidop
4	803	70.7	87286	15	AB026651	Arabidop
5	745	65.6	1104	15	AF373745	Brassic
6	330	34.4	884	15	AF373746	Brassic
7	367.4	32.4	1216	6	AX411603	Sequence
8	338.6	29.8	1159	15	BT013823	Brassic
9	281.6	24.8	12885	14	AF007972	Lotus co
10	256.8	22.6	1444	6	AX411613	Sequence
11	247.4	21.8	1157	15	AY295226	Arabidop
12	237.4	20.9	987	6	AX411601	Sequence
13	221.6	19.5	242	15	AT524543	Arabidop
14	201.6	17.8	1302	6	AX411605	Sequence
15	183.4	16.2	1115	15	AF373544	Brassic
16	163.4	14.4	975	6	AX412268	Sequence
17	163.4	14.4	975	6	AX412469	Sequence
18	163.4	14.4	975	6	AX505463	Sequence

19	163.4	14.4	1139	15	BT000309	Arabidops
20	163.4	14.4	1188	6	E63071	Arabidops
21	163.4	14.4	1188	15	AF021244	Arabidops
22	163.4	14.4	1194	6	AY099815	Arabidops
23	151	13.3	1174	6	AX411607	Sequence
24	148.2	13.1	1242	6	AX411617	Sequence
25	148.2	13.1	1242	15	BT009214	Triticum
26	145.4	12.8	1104	6	AX411609	Sequence
27	142.8	12.6	1155	6	AX411611	Sequence
28	136	12.0	1157	15	AY089091	Arabidops
29	121.2	10.7	992	15	AB056126	Citrus
30	118	10.4	1092	15	AF160869	Citrus si
31	114.4	10.1	9031	15	AC024609	Arabidops
32	114.4	10.1	119842	15	AC007497	Arabidops
33	109	9.6	110000	15	AF008216	Arabidops
34	109	9.6	142737	15	AC027658	Oryza sat
35	109	9.6	103145	15	AB017095	Oryza sat
36	93.6	8.2	1340	6	E63070	DNA encodin
37	93.6	8.2	1340	15	AB025025	Chenopodi
38	93.6	8.2	1340	15	AF134301	Chenopodi
39	88.4	7.8	1156	6	AX411625	Sequence
40	88.4	7.8	1156	15	BT008923	Triticum
41	71.4	6.3	916	6	AX411615	Sequence
42	71.4	6.3	916	15	BT009115	Triticum
43	66.6	5.9	110000	15	AP004274_280	Connatation (281
44	66.6	5.9	162545	15	AP004274	Oryza sat
45	64.8	5.7	1300	6	E63080	DNA encodin

ALIGNMENTS

RESULT 1	
LOCUS	E63072 1135 bp DNA linear PAT 27-AUG-2002
DEFINITION	DNA encoding chlorophyllase and plant transformed by it.
ACCESSION	E63072
VERSION	E63072.1 GI:22553555
KEYWORDS	JP 2001086990-A/3.
SOURCE	Arabidopsis thaliana (thale cress)
ORGANISM	Arabidopsis thaliana
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eustoids II; Brassicales; Brassicaceae; Arabidopsids.
AUTHORS	1 (base 1 to 1135)
TITLE	Tsuchiya,T., Oca,H., Takamiya,K., Harada,S. and Nakat.K.
JOURNAL	DNA encoding chlorophyllase and plant transformed by it Patent: JP 2001086990-A 3 03-APR-2001; KAGOME CO LTD
COMMENT	OS Arabidopsis thaliana (thale cress) PN JP 2001086990-A/3 PD 03-APR-2001 PF 20-SEP-1999 JP 1999266181 PI TORU TSUCHIYA,HIROYUKI OTA,KENICHIRO TAKAMIYA,SATOSHI HARADA, PI KENGO NAKATA PC C12N15/09,A01H5/00,C12N5/10,C12N9/16//C12N9/16,C12R1:19), PC C12N15/00, PC C12N5/00
FEATURES	CC CC CC Key Location/Qualifiers FT CDS (43)..(996). Location/Qualifiers 1..1135 /organism="Arabidopsis thaliana" /mol_type="genomic DNA" /db_xref="taxon:3702"
ORIGIN	
Query Match	100.0%; Score 1135; DB 6; Length 1135;
Best Local Similarity	100.0%; Pred. No.5,7e-271;
Matches 1135, Conservative	0; Mismatches 0; Indels 0; Gaps 0;
1	AAAAAAAAAGTAAAGAAAAGAAAACTAATTAAGAACAAAAAAATGCTCTTTCATCA 60

Db 1 AAAAAAAAAAGAAAAAGAAAAAATTAAGAAACAAAAAAATCTCTCTTCAACA 60
QY 61 AGAAAGCGCTTTGAAGATGGCAATACAAATCAATCTTTAAGCTGATCATCT 120
Db 61 AGAAAGCGCTTTGAAGATGGCAATACAAATCAATCTTTAAGCTGATCATCT 120
QY 121 CGTTGCTGAAAAATTAACACCGCTCTAGAGCTTACACCGCTCGGCAAGAGAGTGTG 180
Db 121 CGTTGCTGAAAAATTAACACCGCTCTAGAGCTTACACCGCTCGGCAAGAGAGTGTG 180
QY 181 GTGGCTACGCGCGTGGAGAAAGAGATTATCCGCTGTATGCTCTCCATGTTACCTT 240
Db 181 GTGGCTACGCGCGTGGAGAAAGAGATTATCCGCTGTATGCTCTCCATGTTACCTT 240
QY 241 CTCTCAAACTCTTCTATCTTCAAGCTTATGTTGCACTGTCTCTTCAAGCTTCACTTC 300
Db 241 CTCTCAAACTCTTCTATCTTCAAGCTTATGTTGCACTGTCTCTTCAAGCTTCACTTC 300
QY 301 ATCGCTCCGATTAATAGTATGCGCGGACCAACACAAATGATGAGATTAAATCAACG 360
Db 301 ATCGCTCCGATTAATAGTATGCGCGGACCAACACAAATGATGAGATTAAATCAACG 360
QY 361 GCGGAGATTATGATTTGTTATCAGTAGACTTAAATCACTTCTTCCAGCGCAATACA 420
Db 361 GCGGAGATTATGATTTGTTATCAGTAGACTTAAATCACTTCTTCCAGCGCAATACA 420
QY 421 CCAAACTATTCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 480
Db 421 CCAAACTATTCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 480
QY 481 GTGCGCTTAAAGAAATTTGGGTAATCTCTGAAATCTTAAAGATCTGCAATGATGCGGTATA 540
Db 481 GTGCGCTTAAAGAAATTTGGGTAATCTCTGAAATCTTAAAGATCTGCAATGATGCGGTATA 540
QY 541 GATCCAGTCGATGAAACAGGAAAGGAAACAAACCCCTCTCGGCTGTGCTTACCTT 600
Db 541 GATCCAGTCGATGAAACAGGAAAGGAAACAAACCCCTCTCGGCTGTGCTTACCTT 600
QY 601 CCAAACTATTCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 660
Db 601 CCAAACTATTCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 660
QY 661 ACCGCTCGGAAACCAATTAATCCACCGTGTGACCTCCCGAGTAGATCACCGAGATTC 720
Db 661 ACCGCTCGGAAACCAATTAATCCACCGTGTGACCTCCCGAGTAGATCACCGAGATTC 720
QY 721 TTTCCGGAATGTCAAGGTCACAGCATGCAATTTGTTGCGAAGATTATGGCATTTGAC 780
Db 721 TTTCCGGAATGTCAAGGTCACAGCATGCAATTTGTTGCGAAGATTATGGCATTTGAC 780
QY 781 ATGCTTGAATGATGATCAAAAGGAGTTAGAGGAAAGATTTCTTATTTGTTGTGAAGAT 840
Db 781 ATGCTTGAATGATGATCAAAAGGAGTTAGAGGAAAGATTTCTTATTTGTTGTGAAGAT 840
QY 841 GGTGAAGAGAGAGACCAATGAGAGATTCGTTGGTGAATCTTGTATCAATTTTGAAG 900
Db 841 GGTGAAGAGAGAGACCAATGAGAGATTCGTTGGTGAATCTTGTATCAATTTTGAAG 900
QY 901 GCTTATTTGAAGAGATGATGATGATTTAGTTAAGATCAAAAGATGGTGTCAAGAGAT 960
Db 901 GCTTATTTGAAGAGATGATGATGATTTAGTTAAGATCAAAAGATGGTGTCAAGAGAT 960
QY 961 GTTCCCGTTGAAGATCAAGAGTTGAGTTATCAAGTAAACATTAAGTTTCTTTAGGGG 1020
Db 961 GTTCCCGTTGAAGATCAAGAGTTGAGTTATCAAGTAAACATTAAGTTTCTTTAGGGG 1020
QY 1021 CTGGTTTTTCTATGTCATATCATCAGCTTTTGTGCTTATGTTTTCATTAAT 1080
Db 1021 CTGGTTTTTCTATGTCATATCATCAGCTTTTGTGCTTATGTTTTCATTAAT 1080
QY 1081 TGTACAACTCTTTAAGTCACTCTTGTCTTCAAAAAAATTTTTTTTTT 1135
Db 1081 TGTACAACTCTTTAAGTCACTCTTGTCTTCAAAAAAATTTTTTTTTT 1135

Db 1081 TGTACAACTCTTTAAGTCACTCTTGTCTTCAAAAAAATTTTTTTTTT 1135
RESULT 2
AF134302 1135 bp mRNA linear PLN 22-JAN-2000
LOCUS Arabidopsis thaliana AtCLH2 (AtCLH2) mRNA, complete cds.
DEFINITION Arabidopsis thaliana AtCLH2 (AtCLH2) mRNA, complete cds.
ACCESSION AF134302
VERSION AF134302.1 GI:6729676
KEYWORDS
SOURCE
ORGANISM
Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 1135)
Teuchlya, T., Ohwa, H., Okawa, K., Iwamatsu, A., Shimada, H., Masuda, T.
and Takamiya, K.
Cloning of chlorophyllase, the key enzyme in chlorophyll
degradation: finding of a lipase motif and the induction by methyl
jasmonate
Proc. Natl. Acad. Sci. U.S.A. 96 (26), 15362-15367 (1999)
10611389
PUBMED
2 (bases 1 to 1135)
Teuchlya, T., Ohwa, H., Okawa, K., Iwamatsu, A., Shimada, H., Masuda, T.
and Takamiya, K.
Direct Submission
Submitted (10-MAR-1999) Department of Biological Sciences, Tokyo
Institute of Technology, Yokohama 226-8501, Japan
location/Qualifiers
source
1. .1135
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
1. .1135
/gene="AtCLH2"
/gene="AtCLH2"
43. .999
/gene="AtCLH2"
/BC_number="3.1.1.14"
/note="chlorophyllase 2; hydrolase; includes a p-loop
motif and lipase/serine active site motif"
/codon_start=1
/product="AtCLH2"
/protein_id="AAP27046.1"
/db_xref="GI:6729677"
/translation="MSSSSSRNAPEDSKYSNLTLDSSRCKITPSSRASPSPKQ
LVAATPVEEDYFVVMHLHGVLVNSFYSLMLHVSHGFLILPQLYSTAGPWTDS
IKSTAEIMDMVSVGLNHFLEPAQVTPNLSKRLSGHSGKTAFAVVALKRGYSNLAI
STLIGIDPVDGKGKQTPPVLAAYLPSNFDLDTPILVIGSGGSETARNDLPFCAP
PGVHREFFRECQSPAHMFVAKDYGHLDMDDPTKGRKSSYCLCKNGEERRRPMRRF
VGGIVSFLKAYLEGGDRELVIKIDGCHDEVPEIQRFEVIM"
ORIGIN
Query Match 100.0%; Score 1135; DB 15; Length 1135;
Best Local Similarity 100.0%; Pred. No. 5,7e-271;
Matches 1135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AAAAAAAAAAGAAAAAATTAAGAAACAAAAAAATCTCTCTTCAACA 60
Db 1 AAAAAAAAAAGAAAAAATTAAGAAACAAAAAAATCTCTCTTCAACA 60
QY 61 AGAAAGCGCTTTGAAGATGGCAATACAAATCAATCTTTAAGCTGATCATCT 120
Db 61 AGAAAGCGCTTTGAAGATGGCAATACAAATCAATCTTTAAGCTGATCATCT 120
QY 121 CGTTGCTGAAAAATTAACACCGCTCTAGAGCTTACACCGCTCGGCAAGAGAGTGTG 180
Db 121 CGTTGCTGAAAAATTAACACCGCTCTAGAGCTTACACCGCTCGGCAAGAGAGTGTG 180
QY 181 GTGGCTACGCGCGTGGAGAAAGAGATTATCCGCTGTATGCTCTCCATGTTACCTT 240
Db 181 GTGGCTACGCGCGTGGAGAAAGAGATTATCCGCTGTATGCTCTCCATGTTACCTT 240
QY 241 CTCTCAAACTCTTCTATCTTCAAGCTTATGTTGCACTGTCTCTTCAAGCTTCACTTC 300
Db 241 CTCTCAAACTCTTCTATCTTCAAGCTTATGTTGCACTGTCTCTTCAAGCTTCACTTC 300
QY 301 ATCGCTCCGATTAATAGTATGCGCGGACCAACACAAATGATGAGATTAAATCAACG 360
Db 301 ATCGCTCCGATTAATAGTATGCGCGGACCAACACAAATGATGAGATTAAATCAACG 360
QY 361 GCGGAGATTATGATTTGTTATCAGTAGACTTAAATCACTTCTTCCAGCGCAATACA 420
Db 361 GCGGAGATTATGATTTGTTATCAGTAGACTTAAATCACTTCTTCCAGCGCAATACA 420
QY 421 CCAAACTATTCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 480
Db 421 CCAAACTATTCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 480
QY 481 GTGCGCTTAAAGAAATTTGGGTAATCTCTGAAATCTTAAAGATCTGCAATGATGCGGTATA 540
Db 481 GTGCGCTTAAAGAAATTTGGGTAATCTCTGAAATCTTAAAGATCTGCAATGATGCGGTATA 540
QY 541 GATCCAGTCGATGAAACAGGAAAGGAAACAAACCCCTCTCGGCTGTGCTTACCTT 600
Db 541 GATCCAGTCGATGAAACAGGAAAGGAAACAAACCCCTCTCGGCTGTGCTTACCTT 600
QY 601 CCAAACTATTCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 660
Db 601 CCAAACTATTCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 660
QY 661 ACCGCTCGGAAACCAATTAATCCACCGTGTGACCTCCCGAGTAGATCACCGAGATTC 720
Db 661 ACCGCTCGGAAACCAATTAATCCACCGTGTGACCTCCCGAGTAGATCACCGAGATTC 720
QY 721 TTTCCGGAATGTCAAGGTCACAGCATGCAATTTGTTGCGAAGATTATGGCATTTGAC 780
Db 721 TTTCCGGAATGTCAAGGTCACAGCATGCAATTTGTTGCGAAGATTATGGCATTTGAC 780
QY 781 ATGCTTGAATGATGATCAAAAGGAGTTAGAGGAAAGATTTCTTATTTGTTGTGAAGAT 840
Db 781 ATGCTTGAATGATGATCAAAAGGAGTTAGAGGAAAGATTTCTTATTTGTTGTGAAGAT 840
QY 841 GGTGAAGAGAGAGACCAATGAGAGATTCGTTGGTGAATCTTGTATCAATTTTGAAG 900
Db 841 GGTGAAGAGAGAGACCAATGAGAGATTCGTTGGTGAATCTTGTATCAATTTTGAAG 900
QY 901 GCTTATTTGAAGAGATGATGATGATTTAGTTAAGATCAAAAGATGGTGTCAAGAGAT 960
Db 901 GCTTATTTGAAGAGATGATGATGATTTAGTTAAGATCAAAAGATGGTGTCAAGAGAT 960
QY 961 GTTCCCGTTGAAGATCAAGAGTTGAGTTATCAAGTAAACATTAAGTTTCTTTAGGGG 1020
Db 961 GTTCCCGTTGAAGATCAAGAGTTGAGTTATCAAGTAAACATTAAGTTTCTTTAGGGG 1020
QY 1021 CTGGTTTTTCTATGTCATATCATCAGCTTTTGTGCTTATGTTTTCATTAAT 1080
Db 1021 CTGGTTTTTCTATGTCATATCATCAGCTTTTGTGCTTATGTTTTCATTAAT 1080
QY 1081 TGTACAACTCTTTAAGTCACTCTTGTCTTCAAAAAAATTTTTTTTTT 1135
Db 1081 TGTACAACTCTTTAAGTCACTCTTGTCTTCAAAAAAATTTTTTTTTT 1135

QY	241	CTCTCAACACTCCTTCTAATCTTCAGCTTATGTTGAGCATGTCCTTCTCAATGAGCTTATCCTC	300
Db	241	CTCTCAACACTCCTTCTAATCTTCAGCTTATGTTGAGCATGTCCTTCTCAATGAGCTTATCCTC	300
QY	301	ATCGCTCCTCAAGTTATATAGTATCGCCGGACCAACATATGATAGATTTAAATCAACG	360
Db	301	ATCGCTCCTCAAGTTATATATAGTATCGCCGGACCAACATATGATAGATTTAAATCAACG	360
QY	361	GCGAGATTAATGGAATTGGTTATCAAGTATAGACTTAAATCATCTTCTTCCAGCCGAACTAACA	420
Db	361	GCGAGATTAATGGAATTGGTTATCAAGTATAGACTTAAATCATCTTCTTCCAGCCGAACTAACA	420
QY	421	CCAAACCTATCCAAATTTTGCCCTCTCCGGCCATACCCCGGGTGGCAAAACCGCGTTTCGG	480
Db	421	CCAAACCTATCCAAATTTTGCCCTCTCCGGCCATACCCCGGGTGGCAAAACCGCGTTTCGG	480
QY	481	GTCGCGCTTAAAGAAATTTGGGGTACTCTCTCGAATCTTAAAGATCTCGACATTTGATCGGTATA	540
Db	481	GTCGCGCTTAAAGAAATTTGGGGTACTCTCTCGAATCTTAAAGATCTCGACATTTGATCGGTATA	540
QY	541	GATCCAGTCGATGAGAACAGGGAAAGGAAACAAACCCCTCCTCCGGTGTGGCTTAACCTT	600
Db	541	GATCCAGTCGATGAGAACAGGGAAAGGAAACAAACCCCTCCTCCGGTGTGGCTTAACCTT	600
QY	601	CCAAACTCATTTGACCTTACGACAAACGCGCTTACTTGTGATCGTTCGGGGCTTGGTGAA	660
Db	601	CCAAACTCATTTGACCTTACGACAAACGCGCTTACTTGTGATCGTTCGGGGCTTGGTGAA	660
QY	661	ACCGGTGGAAACCCATTTATTCGCCACCGGTGACCTCCCGGAGTGAATACACCGAGATGTC	720
Db	661	ACCGGTGGAAACCCATTTATTCGCCACCGGTGACCTCCCGGAGTGAATACACCGAGATGTC	720
QY	721	TTTGGGGAATGTCAAGGTCACAGCATGSGATTTTCGTTGGAGAGATTAATGSGCATTTGGAC	780
Db	721	TTTGGGGAATGTCAAGGTCACAGCATGSGATTTTCGTTGGAGAGATTAATGSGCATTTGGAC	780
QY	781	ATGCTTGAATGATGATACAAAAAGGATTTAGAGGAAAGATTCCTATATTTGTTGTGTAAGAT	840
Db	781	ATGCTTGAATGATGATACAAAAAGGATTTAGAGGAAAGATTCCTATATTTGTTGTGTAAGAT	840
QY	841	GGTGAAGAGAGAGACCAATGAGAGAGATTCGTGGTGACCTTGTGTATCATATTTTGAAG	900
Db	841	GGTGAAGAGAGAGACCAATGAGAGAGATTCGTGGTGACCTTGTGTATCATATTTTGAAG	900
QY	901	GCTTATTTGGAAGAGATGATCGTGAATTAAGTAAATCAAGATGGGTGTCAAGAGAT	960
Db	901	GCTTATTTGGAAGAGATGATCGTGAATTAAGTAAATCAAGATGGGTGTCAAGAGAT	960
QY	961	GTTCCCGGTGAATTCAGAGATTTGAGGTTATCATGTATACATAGATGTTTCTTTAAGGGG	1020
Db	961	GTTCCCGGTGAATTCAGAGATTTGAGGTTATCATGTATACATAGATGTTTCTTTAAGGGG	1020
QY	1021	CTGGTTTTTCTATTGTCAATATCATACAGCTTTTGTGCTTATGTTTACAAACTTATAT	1080
Db	1021	CTGGTTTTTCTATTGTCAATATCATACAGCTTTTGTGCTTATGTTTACAAACTTATAT	1080
QY	1081	TGTACAACCTCTTTAAGTCACTCTTGTCTTACAAAAAATTTTTTTTTTTTTTTTTTTT	1135
Db	1081	TGTACAACCTCTTTAAGTCACTCTTGTCTTACAAAAAATTTTTTTTTTTTTTTTTTTT	1135

RESULT 3	
BT002898	
LOCUS	BT002898 1025 bp mRNA linear PLN 15-JAN-2003
DEFINITION	Arabidopsis thaliana clone RAE15-48-F08 (R20874) putative AtCH2
ACCESSION	BT002898
VERSION	BT002898.1 GI:27754533
KEYWORDS	FLI CDNA.
SOURCE	Arabidopsis thaliana (chale cress)
ORGANISM	Arabidopsis thaliana
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;

REFERENCE	rosid; eucroside II; Brassicales; Brassicaceae; Arabidopsis.
AUTHORS	1 (bases 1 to 1025) Yamada,K., Chan,M.M., Chang,C.H., Dale,J.M., Hsuan,V.W., Lee,J.M., Onodera,C.S., Quach,H.L., Tang,C.C., Toriumi,M., Wong,C., Wu,H.C., Yu,G., Yuan,S., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Kawai,J., Kim,C.J., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shim,P., Southwick,A., Tripp,M.G., Wu,T., Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.
JOURNAL	Arabidopsis Full Length cDNA Clones
COMMENT	2 (bases 1 to 1025) Direct Submission Submitted (15-JAN-2003) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAPL cDNAs (RAPL cDNA : 'RIKEN Arabidopsis Full-length cDNA') : Seki,M., Narusaka,M., Ishida,J., Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J., Hayashizaki,Y. and Shinozaki,K.
FEATURES	The Salk, Stanford, PGEC (SSP) Consortium members carried out the sequencing and annotation of the RAPL cDNAs: Yamada,K., Chan,M.M., Chang,C.H., Dale,J.M., Hsuan,V.W., Lee,J.M., Onodera,C.S., Quach,H.L., Tang,C.C., Toriumi,M., Wong,C., Wu,H.C., Yu,G., Yuan,S., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Kawai,J., Kim,C.J., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shim,P., Southwick,A., Tripp,M.G., Wu,T., Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.
source	Yamada,K. (SSP/PGEC) and Seki,M. (RIKEN GSC) contributed equally to this work. Shinozaki,K. (RIKEN GSC) and Theologis,A. (SSP/PGEC) contributed equally to this work as pis.
Annotation based on July 2002 version of the Arabidopsis genome submitted to Genbank.	
Location/Qualifiers	1..1025 /organism="Arabidopsis thaliana" /mol_type="mRNA" /db_xref="taxon:3702" /chromosome="5" /clone="RAPL15-48-F08 (R20874)" /ecotype="Columbia" /note="This clone is in a modified pBluescript vector2 (lambda PS) as a BamHI/XhoI insert."
gene	<1..1025 /gene="At5g43860"
CDS	<1..883 /gene="At5g43860" /codon_start=2 /evidence=experimental /product="putative AtCh2 protein" /protein_id="AA022714.1" /db_xref="GI:27754534"
misc_difference	5 /gene="At5g43860" /note="compared to genomic sequence resulting in an amino acid sequence difference" /replace="c"
misc_difference	95

```

/gene="At5g43860"
/note="compared to genomic sequence resulting in an amino
acid sequence difference"
/replace="C"
884..1025
/gene="At5g43860"
misc_difference 1010
/gene="At5g43860"
/note="compared to genomic sequence"
/replace="T"

```

ORIGIN

```

Query Match      87.9%; Score 998.2; DB 15; Length 1025;
Best Local Similarity 98.7%; Pred. No: 5,8e-237;
Matches 1006; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Oy 117 ATCTGTTGCTGCAGAAATTAACACCGCTTTCTAGACCTTCACCGTCTCCGCAAGACACT 176
Db 1 ATCTGTTGCTGCAGAAATTAACACCGCTTTCTAGACCTTCACCGTCTCCGCAAGACACT 60

Oy 177 GTTGTGCTAGCGCGGTGAGGAGAGATTAATCCGGTGGATGCTCTCCATGGTTA 236
Db 61 GTTGTGCTAGCGCGGTGAGGAGAGATTAATCCGGTGGATGCTCTCCATGGTTA 120

Oy 237 CCTTCTACAACTCTCTATCTCAGCTTAATGTCATGTCCTCTCTCATGGCTTCAT 296
Db 121 CCTTCTACAACTCTCTATCTCAGCTTAATGTCATGTCCTCTCTCATGGCTTCAT 180

Oy 297 CCTCATGCTCTCTCATGTTATATAGTATGCGCGGACCAACATGATGATGAATTAATC 356
Db 181 CCTCATGCTCTCTCATGTTATATAGTATGCGCGGACCAACATGATGATGAATTAATC 240

Oy 357 AACGGCGGAGATTATGATGGTTATCAGTAGGACTTAATCACTTCTCCAGGCGCAAGT 416
Db 241 AACGGCGGAGATTATGATGGTTATCAGTAGGACTTAATCACTTCTCCAGGCGCAAGT 300

Oy 417 AACACCAAACTATCCAAATTTGCGCTCTCCGCGCATAGCGCGTGGCAAAACCGCGTT 476
Db 301 AACACCAAACTATCCAAATTTGCGCTCTCCGCGCATAGCGCGTGGCAAAACCGCGTT 360

Oy 477 TGGCGTGGCTTTAAAGAAATTTGGTACTCTCGAATCTAAAGATCTCGAATGATGATCGG 536
Db 361 TGGCGTGGCTTTAAAGAAATTTGGTACTCTCGAATCTAAAGATCTCGAATGATGATCGG 420

Oy 537 TATAGATCAGTCGATGGAACAGGGAAGGGAACAAACCCCTCTCCGGGTGGCTTA 596
Db 421 TATAGATCAGTCGATGGAACAGGGAAGGGAACAAACCCCTCTCCGGGTGGCTTA 480

Oy 597 CCTTCCAAACTATTTGACCTAGACAAAACGCTTACTTGTGATCGGTGCGGGCTTGG 656
Db 481 CCTTCCAAACTATTTGACCTAGACAAAACGCTTACTTGTGATCGGTGCGGGCTTGG 540

Oy 657 TGAACCGCTGGAACCCATTAATTTCCACCGTGTGACCTCCGGAAGTAATCACCGAGA 716
Db 541 TGAACCGCTGGAACCCATTAATTTCCACCGTGTGACCTCCGGAAGTAATCACCGAGA 600

Oy 717 GTTCTTTGGGGAATGTCAAGGTCAGACATGGCATTTCGTTGGAGAGATTATGGGCATT 776
Db 601 GTTCTTTGGGGAATGTCAAGGTCAGACATGGCATTTCGTTGGAGAGATTATGGGCATT 660

Oy 777 GACACATGCTTATGATGATACAAAAGGATTAAGAGGAAGAGTTCTTAATGTTGTGTA 836
Db 661 GACACATGCTTATGATGATACAAAAGGATTAAGAGGAAGAGTTCTTAATGTTGTGTA 720

Oy 837 GAATGTGTAAGAGAGAGACCAATGAGAGATTCGTTGGTGAATCTTGTGATCAATTTT 896
Db 721 GAATGTGTAAGAGAGAGACCAATGAGAGATTCGTTGGTGAATCTTGTGATCAATTTT 780

Oy 897 GAAGCTTAATTTGGAAGAGATGATCGTAATTAAGATCAAAAGATGGGTGTCACGA 956
Db 781 GAAGCTTAATTTGGAAGAGATGATCGTAATTAAGATCAAAAGATGGGTGTCACGA 840

Oy 957 GGATGTTCCCGTTGAATTTCAAGAGTTTGAAGTTATCATGTAACATTAAGTTTCTTTA 1016

```

```

Db 841 GGATGTTCCCGTTGAATTTCAAGAGTTTGAAGTTATCATGTAACATTAAGTTTCTTTA 900
Oy 1017 GGGGCTGTTTCTTATGTCATATCATATCATCAGCTTTGTTGCTTATGCTTTACAACTT 1076
Db 901 GGGGCTGTTTCTTATGTCATATCATATCATCAGCTTTGTTGCTTATGCTTTACAACTT 960
Oy 1077 ATATTGACAACTTTTAAGTCACCTCTGCTTACAAAACCAACCAACCAACCAACCA 1135
Db 961 ATATTGACAACTTTTAAGTCACCTCTGCTTATGATATTAACCCGACCAACCAACCA 1019

RESULT 4
AB026651
LOCUS
DEFINITION
AB026651 Ba000015
VERSION
KEYWORDS
SOURCE
ORGANISM
Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
1
Kaneke,T., Katoh,T., Asamizu,E., Sato,S., Nakamura,Y., Kotani,H.
and Tabata,S.
Structural analysis of Arabidopsis thaliana chromosome 5. XI
Unpublished
2 (bases 1 to 87286)
Nakamura,Y.
Direct Submission
Submitted (28-APR-1999) Yasukazu Nakamura, Kazusa DNA Research
Institute, Department of Plant Gene Research, 1532-3, Yana,
Kisarazu, Chiba 292-0812, Japan (E-mail: ynakamu@kazusa.or.jp,
Tel:81-438-52-3935, Fax:81-438-52-3934)
Address for correspondence: kase@kazusa.or.jp
For the latest information on annotation of this clone, please see
http://www.kazusa.or.jp/kaos/cgi-bin/agd_graph.cgi?c=MOD19
Genes with similarity to proteins in the databases are described in
'product' or 'note' qualifiers. Genes that have no significant
protein similarity are described as 'unknown protein'.
The software programs used to predict genes include: Grail
(Informatics Group, Oak Ridge National Laboratory,
http://compbio.ornl.gov/Grail-1.3/),
GENSCAN (Chris Burge, MIT, http://CCR-081.mit.edu/GENSCAN.html),
NetGene2 (S.M. Hebsgaard, et al., CBS, Technical University of
Denmark, http://www.cbs.dtu.dk/services/NetGene2/) and
SplicePredictor (Volker Brendel, Stanford University,
http://gremli.zool.iastate.edu/cgi-bin/ap.cgi).
Genes encoding tRNAs are predicted by tRNAscan-SE
(Sean Eddy, Washington University School of Medicine, St. Louis,
http://genome.wustl.edu/eddy/tRNAscan-SE/).
This sequence may not be the entire insert of this clone. It may be
shorter because we remove overlaps between neighboring submissions.
The 5' clone is MQ024 and the 3' clone is F686.
location/Qualifiers
1..87286
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"
/chromosome="5"
/clone="MOD19"
/clone_1fb="Mitui P1"
/ecotype="Columbia"
1164..2159
/note="gene_id:MOD19.2"
/codon_start=1
/evidence=not experimental
/product="steroid sulfotransferase-like"
/protein_id="BA011286.1"
/db_xref="GI:1017937"
/translation="MAGKSDLPIARDNLSBETKTLISSLPYQDSHVNLCKYQGM"

```

YHNTLAVINVOGNPOBDTIIILASPKSGTWTAKLSVAIYVERSKOPRDDPLTH
PLSDNPHGIYPPPEPMTYKTSPTDLTKSTSSPRLEPTMPLHTEKSGKSPCY
VYMCNKIDVLISDMHFRSKYSNNEVSSTLESNMFESFCGVSFYGPWHDALSVMG
SLENPVKVLFMRVEMKTEPCVOVKRLAEFGPFTKEEDSGSISKLIELCSIGNIS
GLBYNKGCTWMNDYKYSFKRKGSDMKMLTPEMKNKIDMIIEBKLSKSDLKF"
join(5356, .5534, 5627, .5929, 6019, .6097)
/note="gene_id:MOD19.3
ep|p33077"
/codon_start=1
/evidence=not_experimental
/product="auxin-induced protein AUX2-11"
/protein_id="BAB11297.1"
/db_xref="GI:10177938"
/translation="MEKYDVIDELVNLKATRLRGLPETETVSCGSKRKYPEATE
KEISTGTETASPPKAOIVGMPVRSYRKNNVQTKSESSEGGNAYKVSNDGAPYLR
KIDLTMYKQPELMKSLNMFESVFEREGYKSDPVTYDIXOGDMWLVDGVPM
EMFVSCRLIMKGSRYKGLGCGGL"
complement(join(7143, .7430, 7666, .7870, 7958, .8028,
8118, .8253, 8345, .8424, 8530, .8643, 8737, .8818, 8906, .8994,
9168, .9218, 9339, .9448, 9635, .9707, 9969, .10024, 10108, .10198,
10456, .10542, 10681, .10773, 10857, .10976, 11286, .11414))
/note="unnamed protein product; db|BA91806.1
gene_id:MOD19.4
similar to unknown protein"
/codon_start=1
/evidence=not_experimental
/protein_id="BAB11298.1"
/db_xref="GI:10177939"
/translation="MDSNFKWLFLAILISLTPSGFVLHGVLAESVKADEAKOLRDEV
RGMYHAFDGMNNAFLDELRLPSGCEPTLGGVALTILDSLTLLALDGRFESVS
VEWIKGLQFNINKTVSEFTTIRVLGSLASHLASDYANGMIRPENNELLVLAAN
LARMPLPAFTPTGIPPGSVNLMGVNDKHSKTSITGGGLSLSEFVLSLTNDPVE
EQAKNVRGLMARSNLDLVGAHINVFTEBWQKDGISDSFEVYLKYLILG
DEEVLTYFOEAYRSAMQYLNDPMWYVNMDSALIVPFRNSIDAFWPGQVLAGVD
PAIRHTYAFSVMYRKYGTPEGFNATLSVOYQKSYLPELLESTMYLYKATDR
YLDAGDFVSLIOYAKPCGCHITDVEHLKODHMSFPLAETVYKMLPLIAD
SDNLVNDGPYKIFSTEGHLLPTROI SLARHCSYRGYCPNSKTLRQVGEQDS
NDHSNYPHSESPTYGLIKGLCPGITHAKTGFSYLPEKTDREVDNPKVYVTS
SYLISQIVTEKRPQEEEGFTSOSEPTMTLSGSSNDQTELTLESSEDDRSYS
"
complement(join(12391, .12699, 12792, .12855, 12945, .13059,
13177, .13315, 13415, .13542, 13630, .13759, 13977, .14129))
/note="unnamed protein product; gene_id:MOD19.5
unknown protein"
/codon_start=1
/evidence=not_experimental
/protein_id="BAB11299.1"
/db_xref="GI:10177940"
/translation="MAHGVAKRPDSKOTAGSSRSKYLAVEKKKYVSIKQMS
VERPLKQDLPREVRSLVEKLBKQKQDAAHRLAVERKIFPNRKIKPERRKLEKS
IRLKEKQRTSSAGVDVIAEQLSKEDLEYRFPKNEKYSLFTGABDEVLK
RGMKQOIKANIIVAASGKLEBETGSEDGLDLDSDDFDKSSSDADADDELTD
KSAKEAASASGRATSGMSDERRKQNSDRALMPPOARFESRNKLYVQENEMS
SSRNNTSRSESESYNARASANSYSQSSNLSNSDAHHPKRRKPKKKQGVLLT
VELICKQIPE"
15158, .17704
/note="gene_id:MOD19.6"
/codon_start=1
/evidence=not_experimental
/product="disease resistance protein"
/protein_id="BAB11300.1"
/db_xref="GI:10177941"
/translation="MDWLSLIPWNNKIFTAACGFLSDSNYIHMESNLDLOKTEE
LKNGRDLARVSIEDKGLQRLAVNGMUSVOIVSEFDDLEAMSITGLCLFG
YCEDCISSTNYGKWKNELEVKELSKNFEVAAKIIPKAEKHIQTTVGLDTPV
GIAMESLIDELIRGLYGMGIGKTLTLESLNKKFVLESEBPVIVVWVSFOLE
GIODOLIGLRAPKEMERETESKASLNNKRRKPVLLDLMASVDLTKGVPP
SRNGSKIVFTTSKRYCKAMKADKOIKVDCSPDEAMELFRILTVDGIILRSQDIPA
LARIANKCHGLPLALNVIGKAVCKSTVEWRHALNVLSPGKPFGEKMERLPLIK
PSYDSLKNGIKCLFYCSLPEDFEIEKOKLLEYWCBEYINPNRDEGDTQGVDI
IGLVRHLILBCELTDKVMQHDVIREMALMINSDFNQOETICVSGAHVRLIPDI
SWEIVRMSLSTQVERKIASPCNCPNLTLLPNYKLVDSVGFPLMPGLVVDLST
NMSLIELPEEISNLGQYVNLSTLTGSKSLPVGLKTKRLIYINLFTVNLJESLVGIA

TTIPNLQVTLFSLPCVDDITHEBQLKHLKILRATTEDMAMILERVQVDRLASI
RGLCLNMSAPRYILNSVALGSLQOIGIVSCNLSLEIEMDLSEKDRHRTSSPGFQ
LAIITYGLVQDLSMLFQMLKDIQVQSPTEIEIINKQGMSTTKKHRIIVDF
GKLESHTLYOLAEITLCWNYQLPMLRESYVNYCPRLIEDIANFPLKG"
20901, .23489
/note="gene_id:MOD19.7"
/codon_start=1
/evidence=not_experimental
/product="disease resistance protein"
/protein_id="BAB11301.1"
/db_xref="GI:10177942"
/translation="MGQGLIIPWNNKIFTAACGFLSDSNYIHMESNLDLOKTEE
LKNGRDLARVSIEDKGLQRLAVNGMUSVOIVSEFDDLEAMSITGLCLFG
YCEDCISSTNYGKWKNELEVKELSKNFEVAAKIIPKAEKHIQTTVGLDKE
MASSSLMDEIQTGLYGMGIGKTLTLESLNKKFVLESEBPVIVVWVSXDFEG
IQOQILGRILSDKEMERETESKASLNNKRRKPVLLDLMASVDLTKGVPP
RENGSKIVFTTSKRYCKAMKADKOIKVDCSPDEAMELFRILTVDGIILRSQDIPA
LARIANKCHGLPLALNVIGKAVCKSTVEWRHALNVLSAGHPEKMERLPLIKF
SYSLKNGIKCLFYCSLPEDFEIEKOKLLEYWCBEYINPNRDEGDTQGVDI
GLVRAHLILBCELTDNVRKHADVIREMALMINSDFKQOETICVSGAHVRLIPDI
MEIVRTMSFCTQIKIKISCKSPNLSLILDNRLVKSINRPFEMPGVLSLVGIA
NLDLIKLPEEISNLGQYVNLSTLTGSKSLPVGLKTKRLIYINLFTVNLJESLVGIA
ATLPLQVLFKFSYCVVDDILMKELODEHLKILRANVDTVILRIQGDRLASI
RSLCLDMSPTPVILSTIALGSLQOALIMCNIIRIMSEKEREISPTLPSG
SPFKQISTVYINQLSGQDLSWLYAQNKKLSEVCMSPQIEIINKEGMITKLR
DIVPFGNLEDLARQMDLTLCWNYRTLPNRKSINDCPRLPDI FVPLPEKSP
SRPFPP"
complement(join(24326, .24427, 24512, .24622, 24721, .24794,
24880, .24949, 25034, .25081, 25158, .25234, 25321, .25447,
25525, .25617, 25678, .25781, 25883, .26128, 26238, .26304,
26403, .26550, 26640, .26724, 27060, .27140, 27291, .27395,
27483, .27637, 27716, .27814, 27937, .28009, 28264, .28335,
28757, .29010, 29183, .29278))
/note="emb|CA86048.1
gene_id:MOD19.8
strong similarity to unknown protein"
/codon_start=1
/evidence=not_experimental
complement(join(34991, .35221, 35369, .35427, 35507, .35569,
35676, .35733, 35815, .36036))
/note="unnamed protein product; gene_id:MOD19.9
unknown protein"
/codon_start=1
/evidence=not_experimental
/protein_id="BAB11302.1"
/db_xref="GI:10177943"
/translation="MATVILISPKSIPKYTDSKFGARVSDQIVNVYKCGSGRLKIA
KLYSAGLSQIBEDIDEDPIGQETNISIMEDPKYIYGAHRTYVGVOKGFPKGA
ADVIAAVDQNGQGLISCMFLPALMGVFDAPGSLYFGALFTVVCILIMDKPD
QPNFEPQIYKLERGARDKLIINDVNTMSIMDFDKGVWDFTIEDDIATR"
complement(37886, .38544)
/note="gene_id:MOD19.10"
/codon_start=1
/evidence=not_experimental
/product="non-17R retroelement reverse
transcriptase-like"
/protein_id="BAB11303.1"
/db_xref="GI:10177944"
/translation="MGQGLIIPWNNKIFTAACGFLSDSNYIHMESNLDLOKTEE
LKNGRDLARVSIEDKGLQRLAVNGMUSVOIVSEFDDLEAMSITGLCLFG
YCEDCISSTNYGKWKNELEVKELSKNFEVAAKIIPKAEKHIQTTVGLDTPV
GIAMESLIDELIRGLYGMGIGKTLTLESLNKKFVLESEBPVIVVWVSFOLE
GIODOLIGLRAPKEMERETESKASLNNKRRKPVLLDLMASVDLTKGVPP
SRNGSKIVFTTSKRYCKAMKADKOIKVDCSPDEAMELFRILTVDGIILRSQDIPA
LARIANKCHGLPLALNVIGKAVCKSTVEWRHALNVLSPGKPFGEKMERLPLIK
PSYDSLKNGIKCLFYCSLPEDFEIEKOKLLEYWCBEYINPNRDEGDTQGVDI
IGLVRHLILBCELTDKVMQHDVIREMALMINSDFNQOETICVSGAHVRLIPDI
SWEIVRMSLSTQVERKIASPCNCPNLTLLPNYKLVDSVGFPLMPGLVVDLST
NMSLIELPEEISNLGQYVNLSTLTGSKSLPVGLKTKRLIYINLFTVNLJESLVGIA

Query Match 90.7%; Score 803; DB 15; Length 87286;
Best Local Similarity 79.4%; Pred. No. 1,5e-188;
Matches 806; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY	3110	CAGTTATATAGATATGCGCGGACACAGACACATATGATATGATATTAATATCAACGCGGAGATT	369
Db	86386	CAGTTATATAGATATGCGCGGACACAGACACAAATGATATGATTAATATCAACGCGGAGATT	86444
QY	3710	ATGATTTGGTATCAGTAGAGACTTATATCACTTTCTTCCAGCGCAAGTATACACAACTTA	429
Db	86446	ATGATTTGGTATATCAGTAGAGACTTATATCACTTTCTTCCAGCGCAAGTATACACAACTTA	86500
QY	4310	TCCAATTTTGGCTCTCCGGCCATATAGCCGCGGTGGCAAAACCGGCTTTGCGGCTTAA	489
Db	86506	TCCAATTTTGGCTCTCCGGCCATATAGCCGCGGTGGCAAAACCGGCTTTGCGGCTTAA	86566
QY	4910	AAGAAATTTGGGTACTCTCTCCGATTTAAAGATCTTGACATTTGATGATATGATTCAGTC	549
Db	86566	AAGAAATTTGGGTACTCTCTCCGATTTAAAGATCTTGACATTTGATGATATGATTCAGTC	86622
QY	5510	GATGAAACAGGGAAGAGGAAACAAACCCCTCTCCGGTGGTGGCTTACTTCCAACTCA	609
Db	86626	GATGAAACAGGGAAGAGGAAACAAACCCCTCTCCGGTGGTGGCTTACTTCCAACTCA	86686
QY	6110	TTTTCAGCTTACACAAACCGCTTATCTTGATTCGGTTGCGGGCTTTGTGAACCGCTCGG	669
Db	86686	TTTTCAGCTTACACAAACCGCTTATCTTGATTCGGTTGCGGGCTTTGTGAACCGCTCGG	86744
QY	6710	AAACCATTTATTTCCACCGCTGTGACCTCCCGAGTGAATTCACCGAGATTCTTTGGGAA	729
Db	86746	AAACCATTTATTTCCACCGCTGTGACCTCCCGAGTGAATTCACCGAGATTCTTTGGGAA	86800
QY	7310	TGTCAAGGTCCAGCATTTGGCATTTGGTTGCGAAGATTATGGGCAATTTGGACATGCTGAT	789
Db	86806	TGTCAAGGTCCAGCATTTGGCATTTGGTTGCGAAGATTATGGGCAATTTGGACATGCTGAT	86866
QY	7910	GATGATACAAAAGGGATTAGAGGGAAGATTCTTATTTGTTGTGTAGAAATGTGTGAAGAG	849
Db	86866	GATGATACAAAAGGGATTAGAGGGAAGATTCTTATTTGTTGTGTAGAAATGTGTGAAGAG	86922
QY	8510	AGGAGACCAATGAGAGAGATTGGTGGTGGACTTTGTTATCATTTTGAAGCTTATTTTG	909
Db	86926	AGGAGACCAATGAGAGAGATTGGTGGTGGACTTTGTTATCATTTTGAAGCTTATTTTG	86986
QY	9110	GAAAGAGATGATTCGATATTTAGTATTAAGTACAAAGATGGGTGTCACGAGATGTTCCGTT	969
Db	86986	GAAAGAGATGATTCGATATTTAGTATTAAGTACAAAGATGGGTGTCACGAGATGTTCCGTT	87044
QY	9710	GAAATTCAGAGTTTGAAGTTATCATATGAACATATATTTTCTTTAGGGCTGGTTTTT	1029
Db	87046	GAAATTCAGAGTTTGAAGTTATCATATGAACATATATTTTCTTTAGGGCTGGTTTTT	87100
QY	10310	CTATTTGATATATCATCATGACTTTTGTGCTTATATGCTTTTAAACCTTATATTTGTAACAT	1089
Db	87106	CTATTTGATATATCATCATGACTTTTGTGCTTATATGCTTTTAAACCTTATATTTGTAACAT	87166
QY	10910	CTTTAATGACACTCTTTTGCTTATCAAAAATAA	1120
Db	87166	CTTTAATGACACTCTTTTGCTTATGATATTTAA	87196
RESULT 5	AF337545	1104 bp	linear
LOCUS	AF337545	1104 bp	linear
DEFINITION	Brassica oleracea chlorophyllase 2 mRNA, complete cds.		
ACCESSION	AF337545		
VERSION	AF337545.1	GI:24210534	
KEYWORDS			
SOURCE	Brassica oleracea		
ORGANISM	Brassica oleracea		
REFERENCE	Chen, H.-H., Chou, Y.-F. and Shaw, J.-F.		
AUTHORS	Molecular characterization and differential regulation of three		
TITLE	chlorophyllase genes, members of the novel serine esterase gene		

REFERENCE	JOURNAL	family, in broccoli
2	(bases 1 to 1104)	Unpublished
AUTHORS	Chen, H.-H., Chou, Y.-F. and Shaw, J.-F.	
TITLE	Direct Submission	
JOURNAL	Submitted (18-JUN-2001) Institute of Botany, Academia Sinica, Taipei, Taiwan 11529, Republic of China	
FEATURES	Location/Qualifiers	
source	1. 1104	/organism="Brassica oleracea"
		/mol_type="mRNA"
		/culivar="Green king"
		/db_xref="taxon:3712"
		/tissue_type="flower"
		1. 41
5'UTR	42. 1007	/EC_number="3.1.1.14"
CDS		/note="BoC1H2"
		/codon_start=1
		/product="chlorophyllase 2"
		/protein_id="AA051934.1"
		/db_xref="GI:24210535"
		/translation="MSSSSSRNAFVQKPKDLITVDLASRCRCYKTPSSILPPPP
		PKSLIVAPVEBGEHPVVMILHGLTILNYSYSQLMHVSSGPIVIAIPOLYNAGPPT
		IDETKSTAEIIDMLSVGLNHFPLPQVTPNLSKALTGHSRGKTAFAVVALKRGYSS
		LKLSAIGVDPVQSGKQTPPVLIVYENSNLEKMPVLVIGSGEELARPLFPP
		CAPGVNHRFPOBCQPAWHFVAKDGHLDMLDDDTKGRKSSYLCRNGBERKPM
		RRFGIVSFLMAVLEDDDCERLVKIAGCEGVPIEIGFEVVK"
		1008. 1104
3'UTR		
ORIGIN		
Query Match	65.6%; Score 745; DB 15; Length 1104;	
Best Local Similarity	82.5%; Pred. No. 4.9e-174;	
Matches	884; Conservative 0; Mismatches 170; Indels 17; Gaps 2;	
QY	2	AAAAAGTAAAGAAAAAACTAATAAAGACAAAAAAATGTCCTCTTCTTCAATCA 61
DB	1	AGAGAAAAAAAGTAAAAAAATATACAAAGAGAAAAAAATCTCATCTTCTTCAATCA 60
QY	62	GAACGCGCTTGAAGATGGCAATATACAAATCAATCTTAACTTGAGCATCATCTC 121
DB	61	GAACGCGCTTGTGATGGCAATATACAAACAGATCTTTTACAGTGAATTTACATCTC 120
QY	122	GTTCCTGC-----AAATAACACCGTCTTCTAGAGCTTCAACGCTCTCCGCAAGC 172
DB	121	GCTGTGCTGTGCTACAAACGACGCGCATCTTCTTGAGCTCGCGCGCTCTCCAGT 180
QY	173	AGCTGTTGTTGCTACGCGCGGTGAGGAGAGATTAATCCGTGTGATGCTCTCATG 232
DB	181	CGCTTTTGTGGCAACCGCGGTGAGGAGAGATTAATCCGTGTGATGCTCTCCATG 240
QY	233	GTTACCTTCTCAACCTCCCTCAATCTGAGCTTAATGTGCAATGCTCTTCTCATGGCT 292
DB	241	GTTACCTTCTCAACCTCAATCTTATCTCCAGCTTAATGTGCAATGCTCTTCTCATGGCT 300
QY	293	TCATCTCATGCTCTCATGTTATATATAGTATCCGCGGACAGACACATGGATGATTA 352
DB	301	TCATCTCATGCTCTCCGCGATTAATATACATTCGCGGACAGACACATGAGAGATTA 360
QY	353	AATCAACGCGGAGATTAATGATGTTGTTATCATGAGACTTAATCACTTCTTCCAGCGC 412
DB	361	AATCAACGCGGAGATTAATGATGTTGTTATCATGAGACTTAATCACTTCTTCCAGCAC 420
QY	413	AAGTAACACCAAACTTCAAAATTTGSCCTCTCCGCGGCAATGCGCGGCGGAGAAACCG 472
DB	421	AAGTAACACCAAACTTCAAAATTTGSCACTTCAACGCGGCAATGCGCGGCGGAGAAACCG 480
QY	473	CGTTTGGCGTCCCTTAAAGAAATTTGGGATCTCTGAAATCTTAAATCTGACATTTGA 532
DB	481	CATTGCGCGTGGCTTTAAAGAAATTTGGATATCTCATGGAATCTTAAATCTGACATTTGA 540
QY	533	TGGTATATGATCCATGCTGATGAAACAGGAAAGGAAACAAACCTTCTTCCGCTGTTGG 592

Db 541 TCGGTGATGATCCGGTGGATGGACAGGAAAGGTAACAAACCCCTCTCCGGTTTAA 600
 Qy 593 CTTACCTTCCAACTGATTGACCTTAAACAAACGCCCTTATCTTGTATCGGTCGGGC 652
 Db 601 CCTATGAACCAACTGATTAACTAGAAAAGATGCGCTTCTATGTTATGTTCCGGAC 660
 Qy 653 TTGGTGAACCGCTCGGAACCCATTATTCACCCGTGTGACCTCCCGAGATGAATCAC 712
 Db 661 TTGGTGAACCTGCGGAAACCCATTGTTTCCACCGTGTGACCTTACCGAGTGAACACC 720
 Qy 713 GAGAGTCTTTCGGAAATGTCAAGTCCAGCATGGCATTTGTTGCGAAGATTATGAGGC 772
 Db 721 GAGAGTCTTTCAGGAATGTCAAGTCCAGCATGGCATTTGTTGCGAAGATTATGAGGC 780
 Qy 773 ATTGGAATGCTTGTATGATGATTAACAAGGATTAAGAGGAGATTCTTATTTGTTGT 832
 Db 781 ATTGGAATGCTTGTATGATGATTAACAAGGATTAAGAGGAGATTCTTATTTGTTGT 840
 Qy 833 GTAGAGTGTGAAGAGAGAGACCAATGAGAGATTCGTGTGTGACCTTGTATCAT 892
 Db 841 GTAGAGTGTGAAGAGAGAGACCAATGAGAGATTCGTGTGTGACCTTGTATCAT 900
 Qy 893 TTTTGAAGCTTATTTTGAAGAGATGATCGTGAATTAGTTAAGATCAAGATGAGTGC 952
 Db 901 TTTTGAAGCTTATCTGGAAGATGATGATTTGTGAATTGTGAAGATCAAGCTGAGTGC 960
 Qy 953 ACGAGATGTTCCCGTGAATAATCAAGAGTTTGAAGTTATCATGTAACATTAATTTTTC 1012
 Db 961 ATGAAGGTGTTCCGTGAATAATCAAGAGTTTGAAGTTAATAAAGTAA-----TTTTC 1012
 Qy 1013 TTTTGGGGCGTGTCTTCTATGTCATATCATATCATGCTTTGTTGCTTATG 1063
 Db 1013 TTTAGACCTGTTTCTTCTATTTTCAATATTAACAATGATCAACATGCTTATG 1063

RESULT 6
 AF337546 884 bp mRNA linear PLN 22-OCT-2002
 LOCUS AF337546
 DEFINITION *Brassica oleracea* chlorophyllase 3 mRNA, complete cds.
 ACCESSION AF337546.1 GI:24210536
 VERSION
 KEYWORDS
 SOURCE

ORGANISM
Brassica oleracea
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 rosids; euroside II; Brassicales; Brassicaceae; Brassica.
 1 (bases 1 to 884)
 Chen, H.-H., Chou, Y.-F. and Shaw, J.-F.
 Molecular characterization and differential regulation of three
 chlorophyllase genes, members of the novel serine esterase gene
 family, in broccoli
 Unpublished
 2 (bases 1 to 884)
 Chen, H.-H., Chou, Y.-F. and Shaw, J.-F.
 Direct Submission
 Submitted (18-JUN-2001) Institute of Botany, Academia Sinica,
 Taipei, Taiwan 11529, Republic of China
 Location/Qualifiers
 1..884
 /organism="Brassica oleracea"
 /mol_type="mRNA"
 /cultiivar="Green king"
 /db_xref="taxon:3712"
 /tissue_type="floret"
 1..43
 44..685
 /EC_number="3.1.1.14"
 /note="BoC1H3"
 /codon_start=1
 /product="chlorophyllase 3"
 /protein_id="AA51935.1"
 /db_xref="GI:24210537"

REFERENCE
 AUTHORS
 JOURNAL
 TITLE
 JOURNAL
 FEATURES
 source

5'UTR
 CDS

3' UTR

/translation="MSPSFLFFLLIKEMSSSSANSPEDEKXTDLTVGLSSCCW
 KRPSSPTQSPERLLVATPVEGEYPVVMLLHGLVLSNFSQMLHVSIGFTVI
 ALQVSIAGPTDTEITDMLSVGINHFLPQVTPNLSKFLSLSHRSKTAFL
 ALAKKFGVSSDLKISALIGIDVGTFTWNGVQYSGEPFEGDCNDRIVES"
 686..884

Query Match 34.4%; Score 390; DB 15; Length 884;
 Best Local Similarity 85.3%; Pred. No. 8,4e-86;
 Matches 435; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

Qy 34 AAAAAAAAAATGTCCTCTTCTTATCAAGAAAGCCCTTTGAATGGAATGCAATATCA 93
 Db 80 ATAAAGAAATGTCCTCTTCTATCATAGCAAACTCTTGAAGACGCAATACAAACA 139
 Qy 94 AATCTCTTAACCTTGAACATCATCTGTTGTGCAAAATACACGCTTCTAGAGCT 153
 Db 140 GATCTTTTAAAGTAGGCTTATCATCTTGTCTGAAAAAGCCCTCTTCTCGACT 199
 Qy 154 TCACCGTCTCCGCAAGACGCTGTTGTGTGCTACGCGGTGAGAGAGATATCCG 213
 Db 200 CCGCAGTCTCCGCGAAGAGGCTTTGTGTGCAACCGGTGAGAGAGATATCCG 259
 Qy 214 GTGGTATGCTCCATGATGTTACCTTCTTAACTCTTCTATCTAGCTTATGTTG 273
 Db 260 GTGGTATGCTCCATGATGTTACCTTCTTAACTCTTCTATCTAGCTTATGTTG 319
 Qy 274 CATGCTCTTCTATGATGCTTATCTCATGCTCTCATGTTATATATATGCGCGACA 333
 Db 320 CATGCTCTTCTTCCATGATGCTTATCTCATGCTCTCATGTTATATATGCGCGACA 379
 Qy 334 GACACATGATGATGATTAATCAACGCGGAGATTATGATTGTTATCATGAGACTT 393
 Db 380 GACACATGATGATGATTAATCAACGCGGAGATTATGATTGTTATCATGAGACTT 439
 Qy 394 AATCATTCTTCTTCAACGCGAAGTAACACAAACCTTATCCAAATTCCTTCCGCGCAT 453
 Db 440 AATCATTCTTCTTCAACGCGAAGTAACACAAACCTTATCCAAATTCCTTCCGCGCAT 499
 Qy 454 AGCCGCGGTGCAAAACCGGTTTGGGCTGCGCTTAAAGAAATTTGGTACTCTCGAAT 513
 Db 500 AGCCGCGGTGCAAAACCGGTTTGGGCTGCGCTTAAAGAAATTTGGTACTCTCGAAT 559
 Qy 514 CTAAAGATCTCGACATTGATCGGTATAGAT 543
 Db 560 CTAAAGATCTCGGCAATTGATAGGTATAGAT 589

RESULT 7
 AX411603 1216 bp DNA linear PAT 14-JUN-2002
 LOCUS AX411603
 DEFINITION Sequence 3 from Patent WO0229022.
 ACCESSION AX411603
 VERSION AX411603.1 GI:21444160
 KEYWORDS
 SOURCE
 ORGANISM
Vitis sp.
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 rosids; Vitaceae; Vitis.
 1
 Thorpe, C., Cahoon, E.B. and Cahoon, R.E.
 Chlorophyllases
 Patent: WO 0229022-A 3 11-APR-2002;
 R.I. DU PONT DE NEMOURS AND COMPANY (US)
 Location/Qualifiers
 1..1216
 /organism="Vitis sp."
 /mol_type="unassigned DNA"
 /db_xref="taxon:3604"

REFERENCE
 AUTHORS
 JOURNAL
 TITLE
 JOURNAL
 FEATURES
 source

5'UTR
 CDS

ORIGIN

Query Match 32.4%; Score 367.4; DB 6; Length 1216;

Best Local Similarity 63.9%; Pred. No. 3.4e-80;			
Matches 608; Conservative 0; Mismatches 331; Indels 12; Gaps 3;			
Qy	50	CTTCTTCATCAAGAAAGCCCTTTGAAGTGGCAATATCAATATCTTTAACCTTGG	109
Db	1	CTTCTTCATGTTTCCACCACTGCTGCAAAATGTTTTGAATTTGAAAGCAATCAAGTGC	60
Qy	110	ACTCATCTCTGCTGTCGCAAAATPAACCGCTTCTTAGAGCTTCAACGCTCCGCCAA	169
Db	61	TTCTTAGAGCTGAACCAAGCACTTTCACACCAAGTCACTCTTCTTCCCTCCAC	120
Qy	170	AGCAGCTTGTGTGCTACGCGGCTGAGAGAAAGATTAATCCGCTGTGATGCTCTCC	229
Db	121	TGCAACTCTGTATGCTACGCGCTGAGAGAGGGAGTCCCGCTGCTGCTCTCCCTC	180
Qy	230	ATGCTTACCTTCTCAACAATCTCTTCTATCTCAGCTTATGTTGATGATCTCTCTCATG	289
Db	181	ATGCTTATCTCTCTATTAATCTCTTCTATCTCCAGCTTATCCACATAGCTCTCTCATG	240
Qy	290	GCTTCATCTCATGCTCTCATATATATGATGCGCGAGCCAGACAAATGATGAGA	349
Db	241	GTTTCATGTTCTTCTCTCTCTCATATATCACTGTGCTGACCAATTCAGCGAAGCA	300
Qy	350	TTAAATCAACGCGGAGATTAATGATGCTTATCAGTAGAAGTTATCATCTTCTTCCAG	409
Db	301	TCAAGTCCGAGCTCTTTAACAAAATGCTTATCCAAAGAGCTCATGATCTTACTCTC	360
Qy	410	CGCAGTAACACCAAACTATATCCAAATTTGCCCTTCCGCGCATATGCCGCTGCAAAA	469
Db	361	CCCAATGTCGCGCAAAATTTAAAGCAATGAGCATTTGCGCGCATATGCTGAGAGCAAAA	420
Qy	470	CCGCTTGGCGGCTCTTAAAGAAATTTGGGTAATCTCTCGATCTTAAAGATCTGCAAT	529
Db	421	CTGCTTTTGTCTAGACTGAGAAA-----GCATCACTTCTCTGAAATTTTCAGCT	474
Qy	530	TGATCGTATAGATCACTGATGAGACAGGAAAGGAAACAAACCCCTCTCCGCTGT	589
Db	475	TGATAGGATAGACCCGCTGATGAGATGACAAAGGAAACAAACCCCTCCACCGGATAC	534
Qy	590	TGGCTTACCTTCCAACTCATTTGACTAGACAAACCCCTATCTTGTGATCGCTTGG	649
Db	535	TCACCTATGTTCTCATTTCAATTTGATCTAGAC--ATGCAAGTGAATGGTAAATGGTTCGG	591
Qy	650	GACTTGTGAAACCGCTCGGAACCCATATTTCCACCGTGTGACCTCCCGAGTGAATC	709
Db	592	GTTTGGGTGAAGTGAAGAAACCTCTGTCTCTCTGCTGCTCCCAAGGGCGTAAAC	651
Qy	710	ACCGAGCTTCTTGGGAATGTCAAGGTCCAGATGCGATTTCTTGGCGAAGATTAAG	769
Db	652	ATGAGGACTTCTTAAAGAAATGCGGTGAACAGCTTGTATTTCTTSCAAAGACTATG	711
Qy	770	GGCATTGTGACATGCTTATGATATGATACAAAGGATTAAGAGGAAAGTCTTATCTT	829
Db	712	GCCATCTTGACATGCTAATGATGACTAATGAGATTAAGAGGAAAGCTTACACTTGT	771
Qy	830	TGTGTAAAGATGTGAAGAGAGAGACCAATGAGAGATTCGTTGTGATCTTGTAT	889
Db	772	TGTGTAAAGATGTGAAG	831
Qy	890	CATTTTGAAGCTTATTTGAAGAGATGATCTGAATTAAGTTAAGTCAAAAGATGGT	949
Db	832	CATTTTGAAGCTTATTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	890
Qy	950	GTCACAGAGATTTCCGCTTGAATTTCAAGATTTGAAGTATCATGTAA	1000
Db	891	--CATGTACTGCAACCAAGTGAAGCTTCAAAATGTTGAGTTCCTCGTGA	939

RESULT 8
LOCUS BT013523 1159 bp mRNA linear PLN 11-MAY-2004
DEFINITION Lycopersicon esculentum clone 132241R, mRNA sequence.
ACCESSION BT013523

VERSION BT013523.1 GI:47104938			
KEYWORDS FLI_CDNA			
SOURCE Lycopersicon esculentum (Solanum Lycopersicum)			
ORGANISM Lycopersicon esculentum			
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.			
AUTHORS 1 (bases 1 to 1159)			
TITLE Kirchner, E.F., Wang, W. and Vazelle, A.			
JOURNAL Direct Submission			
FEATURES Submitted (11-MAY-2004) The Institute for Genomic Research, 9712			
source Location/Qualifiers			
1. 1159			
/organism="Lycopersicon esculentum"			
/mol_type="mRNA"			
/db_xref="taxon:4081"			
/clone="132241R"			
/issue_type="fruit"			
/note="TXBH55"			
ORIGIN			
Query Match 29.8%; Score 338.6; DB 15; Length 1159;			
Best Local Similarity 64.0%; Pred. No. 5e-73;			
Matches 547; Conservative 0; Mismatches 299; Indels 9; Gaps 2;			
Qy	160	TTCTCCGCAAGCAGCTGTTGCTGCTACGCGCTGAGAGAGATTAATCCGCTGCTG	219
Db	149	TTCTCTCAAAATCTCTGTTAATTTGGACCCATCAAGACAGAAATCTCCAGATTTG	208
Qy	220	ATGCTCTCCATGTTTACTTCTCTACAACTCTTCTATTTCTGAGTTATGTTGATGTC	279
Db	209	ATATTTCTCATGCTATCTTCTTCAATTTCTTCAATCAACATTAATCAACATCTC	268
Qy	280	TCTTCTATGCTTATCTCATGCTCTCTCATGTTATATATATGCTGCGGACAGACA	339
Db	269	TCTTCTATGCTTATGTTGTTGCTCTCATGTTATATATATGTTGTTGAGAGAGATGCA	328
Qy	340	ATGATGATTAATTAATCAAGGCGGAGATTATGATGTTATTCAGTAGAGACTTAATCAC	399
Db	329	ACCGAGATTAATTAATGACAGCTGAGTCAACCATTTGTTATCTGAAGATTCACAT	388
Qy	400	TTTCTTCACGCAAGTATACCAAACTATCCAAATTTGCCCTTCCGCGCATAGCCGC	459
Db	389	CACCTTCACTGATATGTTGAGCCAACTTGACAACTTGAGACTGAGCGCATAGCGCT	448
Qy	460	GTTGCGCAAAACCGCTTGGGCTGCTTAAAGAAATTTGGGTTACTCTGAAATCTAAG	519
Db	449	GGGGGAAGTTGATTTTGGCTAGCTCTTGTAGACTTG-----CTAGTGAACCTGAAA	502
Qy	520	ATCTGCACTTGAATCGGTATATGATCCAGTATGAGAAACAGGAAAGGAAACAAACCCCT	579
Db	503	TTTTACGCTTATGATGCTGTTGATCTCTGTGATGAAATGAGAAAGGAAAGCAATCTCG	562
Qy	580	CTTCCGCTGTTGCTTCACTTCAAACTATTTGACCTAGACAAACGCTTATCTTGTG	639
Db	563	CCATCACTTCACTTCACTTCCGCTTCTTCAATTAATCTTGTATATGCTGATGATA	622
Qy	640	ATCGCTTGGGCTTGTGTAACCGCTCGGAACCATTAATTTCCACCGTGTGACCTCCC	699
Db	623	ATTGCTCGGGTTTGGAGAAATTAAMAAAGATCTCTGTTCCCTGCTTGGCTCTTAAA	682
Qy	700	GGAGTAATCACCGAGCTTCTTCCGGAATGTCAAGTCCAGATGCGCAATTTGTTGGG	759
Db	683	GGGAGTAATCACCGATTTCTACACGAATGTGCAAGCAGCTTCTTACTTGTGTGCA	742
Qy	760	AAAGATTATGAGCTTTGAGATGCTTATGATGATCAAAAGGATTAAGGAGAGAGT	819
Db	743	AAAGATTATGAGCAATATGATATGATGATGAGAGACAAAGGATTTGAGGAGAGCT	802
Qy	820	TCTTATTTGTTGTGAAGATGTTGAAGAGAGAGACCAATGAGAGATTTGTTGTGA	879
Db	803	ACTTATTTGTTGTGCAAGAAAGGAAATCTTAGAGAGCCATGAGAGATTTGTTGAGG	862

QY 880 CTGTGATCATTTTGAAGCTTATTTGAGAGATGATCGTAATTGTTAAGTAC 939
DB 863 GTTTTGCGCATTTTGGAGCTTATTTGAGAGATTTCCAGCGCTTATAGCTATT 922
QY 940 AAGATGGGATCTACGAGATGTTCCCGTGAATAATCAAGTTGAGGTATCATGTA 999
DB 923 AGAATG---GCCATGTTGCACCTGCTGTAGAGCTCCAGATATTGATTTCTGTCTAA 979
QY 1000 ACATATGTTTTCCTT 1014
DB 980 GCTATGTTGTGCTT 994

RESULT 9
AP007972 128855 bp DNA linear HTG 28-DEC-2004
LOCUS locus corniculatus var. japonicus chromosome 5 clone LJt44A18, ***
DEFINITION SEQUENCING IN PROGRESS ***, 45 unordered pieces.
ACCESSION AP007972.1 GI:56806280
VERSION AP007972.1
KEYWORDS HTG, HTGS, PHASE1.
SOURCE locus corniculatus var. japonicus (Lotus japonicus)
ORGANISM Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Lotaeae; Lotus.

REFERENCE 1
AUTHORS Kaneke, T., Aeamizu, E., Nakamura, Y., Sato, S. and Tabata, S.
TITLE Structural Analysis of a Lotus japonicus Genome. XI. Sequence Features and Mapping of Nine hundred twenty-one TAC Clones
JOURNAL Unpublished
AUTHORS Sato, S.
REFERENCE 2 (bases 1 to 128855)
JOURNAL Direct Submission
COMMENT Submitted (26-OCT-2004) Shusui Sato, Kazusa DNA Research Institute, Department of Plant Gene Research; 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 252-0818, Japan (E-mail:ssato@kazusa.or.jp), URL: http://www.kazusa.or.jp/, Tel: 81-438-52-3935 (ex. 2337), Fax: 81-438-52-3934)
* NOTE: This is a 'working draft' sequence. It currently consists of 45 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 1715: contig of 1715 bp in length
1716 1815: gap of unknown length
1816 2641: contig of 826 bp in length
2642 2741: gap of unknown length
2742 3555: contig of 814 bp in length
3556 4548: gap of unknown length
4549 4648: gap of unknown length
4649 6180: contig of 1532 bp in length
6181 7099: contig of 819 bp in length
7100 7199: gap of unknown length
7200 8339: contig of 1140 bp in length
8340 8439: gap of unknown length
8440 9486: contig of 1047 bp in length
9487 9586: gap of unknown length
9587 11121: contig of 1535 bp in length
11122 11221: gap of unknown length
11222 12794: contig of 1573 bp in length
12795 12894: gap of unknown length
12895 14686: contig of 1792 bp in length
14687 16382: contig of 1596 bp in length
16383 16482: gap of unknown length

FEATURES

Source

16483 17841: contig of 1359 bp in length
17842 17941: gap of unknown length
17942 19137: contig of 1196 bp in length
19138 19237: gap of unknown length
19238 21270: contig of 2033 bp in length
21271 21370: gap of unknown length
21371 22827: contig of 1457 bp in length
22828 22927: gap of unknown length
22928 24839: contig of 1912 bp in length
24840 27581: gap of unknown length
27582 27681: gap of unknown length
27682 29435: gap of 1754 bp in length
29436 29535: gap of unknown length
29536 31538: contig of 2003 bp in length
31539 31638: gap of unknown length
31639 33866: contig of 2228 bp in length
33867 33966: gap of unknown length
33967 36055: contig of 2089 bp in length
36056 36155: gap of unknown length
36156 38250: contig of 2095 bp in length
38251 38350: gap of unknown length
38351 40933: contig of 2583 bp in length
40934 41033: gap of unknown length
41034 43319: contig of 2286 bp in length
43320 43419: gap of unknown length
43420 45644: contig of 2225 bp in length
45645 45744: gap of unknown length
45745 47854: contig of 2110 bp in length
47855 47954: gap of unknown length
47955 50383: contig of 2429 bp in length
50384 50483: gap of unknown length
50484 53820: contig of 3337 bp in length
53821 53920: gap of unknown length
53921 56225: contig of 2305 bp in length
56226 56325: gap of unknown length
56326 59890: contig of 3565 bp in length
59891 59990: gap of unknown length
59991 62614: contig of 2624 bp in length
62615 62714: gap of unknown length
62715 65635: contig of 2921 bp in length
65636 65735: gap of unknown length
65736 69840: contig of 4105 bp in length
69841 69940: gap of unknown length
69941 73812: contig of 3872 bp in length
73813 73912: gap of unknown length
73913 79015: contig of 5103 bp in length
79016 79115: gap of unknown length
79116 83508: contig of 4393 bp in length
83509 83608: gap of unknown length
83609 87605: contig of 3997 bp in length
87606 87705: gap of unknown length
87706 92715: contig of 5010 bp in length
92716 92815: gap of unknown length
92816 98011: contig of 5196 bp in length
98012 98111: gap of unknown length
98112 102876: contig of 4765 bp in length
102877 102976: gap of unknown length
102977 108287: contig of 5311 bp in length
108288 108387: gap of unknown length
108388 112790: contig of 4403 bp in length
112791 112890: gap of unknown length
112891 127873: contig of 14983 bp in length
127874 127973: gap of unknown length
127974 128855: contig of 882 bp in length.

location/Qualifiers
1. 128855
/organism="Lotus corniculatus var. japonicus"
/mol_type="genomic DNA"
/variety="japonicus"
/db_xref="taxon:34305"
/chromosome="5"
/clone="LJt44A18"

Matches 474; Conservative 0; Mismatches 287; Indels 15; Gaps 3;

QY 154 TCACCTCTCCGCCAAGACCTGTGGCTACGCCGGTGAAGAGAGATTCG 213
 199 TCCACACACCTCCGAGGCCACACGACGACCACTCCATCTTGAGGAGCGAGTACCA 258
 QY 214 GTGGTANAGCTCCCAAGGTGTACCTTCTCAACAACCTCTTATTTCTAGCTATG 273
 259 ACGCTTCTACTCTTCAAGATGTAGTGTCTCAACAACCTTCTTACTGAGCTTATCAG 318
 QY 274 CATGCTCTTCTCAAGCTTCACTTCATCTCATCTCTCACTTATATATAGTATCGCGAGCA 333
 319 CAATGCGATCCCATGGCTTATATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 378
 QY 334 GACCAATGATGATGATTAATCAACGCGGAGATTAAGATTGTTATCATGATGAGATT 393
 379 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 438
 QY 394 AATGATCTTCTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 453
 439 CAGATGATCTTCCCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 498
 QY 454 AGCCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 513
 499 AGCGGTGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 552
 QY 514 CTAAAGATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 573
 553 TTAAAGCTTCTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 612
 QY 574 ACCCTCTCCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 633
 613 ACCAATCTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 669
 QY 634 CTGTGATCGGCTTCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 693
 670 TTAGTATCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 723
 QY 694 CTCTCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 753
 724 CTAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 783
 QY 754 GTTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 813
 784 GTTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 843
 QY 814 AAGAGTCTTATGTTGT 873
 844 ATTGTGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTC 903
 QY 874 GTTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 929
 904 GTTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 959

RESULT 11
 AY292526 1157 bp mRNA linear PLN 05-JUN-2003
 LOCUS Ginkgo biloba chlorophyllase (CLH) mRNA, complete cds.
 DEFINITION AY292526
 ACCESSION AY292526
 VERSION AY292526.1 GI:31415499
 KEYWORDS
 SOURCE Ginkgo biloba (maidenhair tree)
 ORGANISM Ginkgo biloba
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Ginkgoales; Ginkgoaceae; Ginkgo.
 1 (bases 1 to 1157)
 AUTHORS Okazawa, A., Tang, L., Fukusaki, E.-I. and Kobayashi, A.
 TITLE Direct Submission
 JOURNAL Submitted (06-MAY-2003) Department of Biotechnology, Osaka
 University, 2-1, Yamadaoka, Suita, Osaka 565-0871, Japan
 FEATURES
 source 1..1157

gene
 CDS
 /organism="Ginkgo biloba"
 /mol_type="mRNA"
 /db_xref="taxon:3311"
 1..1157
 /gene="CLH"
 /note="GBCLH"
 55..1083
 /gene="CLH"
 /note="GBCLH"
 /product="chlorophyllase"
 /protein_id="AAP44978.1"
 /db_xref="GI:31415500"
 /translation="MVLVDVSESGPLPVQIILPQANSPCKLADKNGTATTPSPC
 RPKPLILALPSQGDYPLILPHGYVLNLSFQSLRLVASHGYLIAIPOMYSVIG
 NTPETIADAAITDMLRDLSDILPOLNNHVPNPEKLVLAGSHRGKVAFALAGR
 VSOPLKYSALVLDLPVDGKQOQTSHPILSYRESPLDGMGTLVVGSLGCKNRP
 LRPKCAPOGVNHHDPFYBCVAPVHVASDYGLDLDDDTGIGIRKATYCLKNGEA
 REMRKFSSGIVVAFILQAFILGNRGLNDIMVPSHAPVKIEPESLVTEDEVSPVE
 LNRRAVCR"

ORIGIN
 Query Match 21.8%; Score 247.4; DB 15; Length 1157;
 Best Local Similarity 59.9%; Pred. No. 2.3e-50;
 Matches 476; Conservative 0; Mismatches 301; Indels 18; Gaps 3;

QY 163 CGGCCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 222
 190 CTCTCAAAACCTGCTGATGCT 249
 QY 223 CTCTCTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 282
 250 TTTTTCACGCGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 309
 QY 283 TCTCATGAGCTTCACTCTCATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 342
 310 TCCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 369
 QY 343 GAT 402
 370 CCAAGAAATAGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 429
 QY 403 CTTC-----AGCGAAGTAAACCAAACTCTTCAATTTGGCTCTCCGC 450
 430 CTTCGCAAGCTTAAACAATCATGAGGCCCAATTTGAGAAATTTGTGTAGGGGG 489
 QY 451 CATAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 510
 490 CACTCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 548
 QY 511 AATCTAAAGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 570
 549 --TTTAAAGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 606
 QY 571 GAAACCCCTCTCCGAGT 630
 607 GAAACCAATCATCTTCTCTCATGAGAGCAATTTCTTGAATTTG--GTATGCCA 663
 QY 631 ATACTGTATCGGCTTCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 690
 664 ACATTTGT 723
 QY 691 GCACTCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 750
 724 GCTCCCAAGAGGTGTAACCACTGATTTCTTCTCAAGATGATGATGATGATGATGATGATGAT 783
 QY 751 TTGCTGCAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 810
 784 TTTGTGCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 843
 QY 811 GGAAGAGTCTTATGTTGT 870
 844 GGAAGGCTACTTATGCTCTGTATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 913

QY 871 GTTGAGTGAAGCTTGTATCAATTTTGAAGCTTATTTGAAGAGATGATCGTGAATTA 930
 DB 904 AGCGGTGAAGATTTGGTGTGATTTCTTCAAGCATTTTGGTATATATCGAGAGCCCTG 963
 QY 931 GTTAGATCAAGAT 945
 DB 964 AATGATATTAATGTT 978

RESULT 12
 AA411601
 LOCUS AA411601 987 bp DNA linear PAT 14-JUN-2002
 DEFINITION Sequence 1 from Patent WO0229022.
 ACCESSION AX411601
 VERSION AX411601.1 GI:21444159
 KEYWORDS
 SOURCE Vitis sp.
 ORGANISM Vitis sp.
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 rosids; Vitaceae; Vitis.

REFERENCE 1
 Thorpe, C., Cahoon, E. B. and Cahoon, R. E.
 TITLE Chlorophyllases
 JOURNAL Patent: WO 0229022-A 1 11-APR-2002;
 E.I. DU PONT DE NEMOURS AND COMPANY (US)
 FEATURES location/Qualifiers
 source 1..987
 /organism="Vitis sp."
 /mol_type="unassigned DNA"
 /db_xref="taxon:3604"

ORIGIN

Query Match 20.9%; Score 237.4; DB 6; Length 987;
 Best Local Similarity 56.3%; Pred. No. 7.1e-48;
 Matches 551; Conservative 0; Mismatches 411; Indels 17; Gaps 5;

QY 158 GGTCTCGGCAAGAGCGTTGGTGGCTAGCGCGGTGAGAGAGATATATCCGGTGG 217
 DB 23 CTTCCTCCCAAGCATTTGATTTGTTACCAACAATTCACAGGACATACCCAGTTC 82
 QY 218 TGATGCTCTCCATGATGATCTTCTCTCAACTCCCTCTATTTCTGAGCTTAATGTTGATG 277
 DB 83 TCTGTTTCTTCATGCTTTCGAGCTCCGACACACTTCTACACTGAGTCTCTCAACTCA 142
 QY 278 TCTCTTCTCATGCTTCACTCTCATGCTCTCTCACTTAATATATATATATATATATATAT 337
 DB 143 TTTCTTCCCATGATTCATTTGGTGGCTCTCCATGATTAAGGACTATTAATCTCTCTG 202
 QY 338 CAATGATGATTAATCAACGCGGAGATTAAGATGATGATGATGATGATGATGATGATGATGAT 397
 DB 203 GATATCAAGATTAATCAACGAG 262
 QY 398 ACTTCTTCAAGCGCAAGTAACACCAAACTATCCAAATTTGCCCTCTCCGAGCATAGCC 457
 DB 263 CTGTCTCCCAAGAAATGGAACCAAGACTTCAAGCTTGTCTTTCAAGGACCAAGCA 322
 QY 458 GCGGTGCAAAACCGGTTTGGGCTGCTTAAGAAATTTGGGATATCTCTCGAATCTTA 517
 DB 323 GAGGGGGAAGACAGCATTTGTCTGGGCACTAGGATATGATGATCATCTC-----CTCA 376
 QY 518 AGATCTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 577
 DB 377 ACTTCTGAGCCCTACTAGACTAGACCTGTTGGGTTGAGTAATGTTGCCAAAG 436
 QY 578 CTCCTCCGCTGTGGCTTACCTTCCAACTATTTGACTAGACAAAGCGCTTAATCTTG 637
 DB 437 TTCCCAAAATCTAATCTATGTTCTCATTTCTTCAATCTAG---CAATCCAGTTTTCG 493
 QY 638 TGATGCTTGGGCTTGGTGAACCGCTCGAACCCTATATTTCCACCGTGTGACCTTC 697
 DB 494 TAAATCGGACGCGGCTTGGGCGATGAGCAAGGAATCGCTTAACATGATCATATGTCCTCAG 553

QY 698 CCGAGTGAATCACCAGAGTCTTTCGGGAATGTCAAGTCCAGCATGAGCATTTCTTG 757
 DB 554 ATGAGTGAACCATGATGATGATTTTCAAGTGTAAACCTCTGTTCCACTTTGGA 613
 QY 758 CGAAGATTAATGGGATTTGGACATGCTTGATGATATACAAAGGATTTAGAGGAGA 817
 DB 614 CTACTGAATATGTCACCTTGACATGATATGATATCTTTCAAGCTGACATTTGGGGGA 673
 QY 818 GTTCTTATGTTTGTGAAGATGTGAAG---AGAGAGACCAATGAGAGATTCGTTG 874
 DB 674 TTTGCGGTTATATCTGCAAGGTGGAAGGCTCTTAGGACCCCATGAGGAGATGTGG 733
 QY 875 GTGACCTTGTGATCAATTTTGAAGCTTATTTGAAGAGAGATGATGATGATTAAGTTA 934
 DB 734 GTGGCTTTTGTGTGCAATCTTGAAGCTTATTTGAAGGTCACATCTGAGATTT---TCA 790
 QY 935 AGATCAAGATGGGTGTACAGAGATGTTCCCTTGAATTCAGAGATTTGAGGTTATCA 994
 DB 791 AAGCCATTTGATGAACCTGATCTGCTCTGTAAGCTTGATCTGTTGATGATCANAG 850
 QY 995 TGTAACATTAAGTTTCTTTAGGCGCTGTTTCTTATGTCATATCATACGCTTTG 1054
 DB 851 AAGCATGAATCATGAGATCTTTGTAATGCTATAGAGAGGTGA--CAGCATTTT 908
 QY 1055 TTGCTTATGTTTACAACTTATATGTAACAATCTTTAAGTACACCTCTTGTGTTCAA 1114
 DB 909 TTGCTTCTTCTTTGGGAACAAATTCATTAATACCAAACTGTTGTTATTTGCTA 968
 QY 1115 AAAAAAAAAAAAAAAAAA 1133
 DB 969 AAAAAAAAAAAAAAAAAA 987

RESULT 13
 ATH524543
 LOCUS ATH524543 242 bp DNA linear PLN 29-MAR-2003
 DEFINITION Arabidopsis thaliana T-DNA flanking sequence, left border, clone 076H05.
 ACCESSION AJ524543
 VERSION AJ524543.1 GI:26792779
 KEYWORDS left border; T-DNA flanking sequence.
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 rosids; eurosid II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE 1
 Brunaud, V., Balzerque, S., Dubreucq, B., Aubourg, S., Samson, F.,
 Chauvin, S., Bechtold, N., Cruaud, C., Dehose, R., Pelleter, G.,
 Lepoint, L., Caboche, M. and Lecharny, A.
 TITLE T-DNA integration into the Arabidopsis genome depends on sequences
 of pre-insertion sites
 JOURNAL EMBO Rep. 3 (12), 1152-1157 (2002)
 PUBMED 12446565
 REFERENCE 2 (bases 1 to 242)
 AUTHORS Balzerque, S.
 TITLE Direct Submision
 JOURNAL Submitted (21-NOV-2002) Balzerque S., UMRGV, INRA/CNRS, 2 rue
 Gaston Crémieux, 91057 Evry cedex, FRANCE
 COMMENT PCR was performed on DNA from transformants of Arabidopsis thaliana
 plants from INRA (Versailles). The DNA fragment(s) resulting from
 the PCR were directly sequenced from the left or the right border
 to determine the genomic sequence flanking the insertion. T-DNA
 derived sequences were removed. Information to order the
 corresponding mutant line and a link to a database providing a
 graphical display of the insertion site are available at
 http://dbsgap.versailles.inra.fr/publiclines/. This sequence has
 been generated in the framework of the French plant genomics
 program 'Genoplante' (http://www.genoplante.com and
 http://genoplante-info.inbio.gen.fr).
 location/Qualifiers
 source 1..242

```
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"
/clone="076H05"
/clone_1lb="Arabidopsis thaliana T-DNA insertion lines"
/ecotype="Maslilewskija"
1..242
/misc_feature
/feature="T-DNA flanking sequence
left border"

ORIGIN
Query Match 19.5%; Score 221.6; DB 15; Length 242;
Best Local Similarity 98.2%; Pred. No. 6.4e-44;
Matches 224; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 87 CAAATCAATCTTTACCTTGATCATCTGTTGCTGCAAAATPACACCGCTTTC 146
DB 1 CAAATCAATCTTTACCTTGATCATCTGTTGCTGCAAAATPACACCGCTTTC 60
QY 147 TAGAGCTTACCGTCTCCGCCAAGACAGCTGTGTGCTACGCCGGTGAAGAGAGA 206
DB 61 TCGGACTTACCGTCTCCGCCAAGACAGCTGTGTGCTACGCCGGTGAAGAGAGA 120
QY 207 TTATCCGGTGTGATGCTCTCATGCTTCACTTCTCTAGCAACTCTTCTATCTCAAGCT 266
DB 121 TTATCCGGTGTGATGCTCTCATGCTTCACTTCTCTAGCAACTCTTCTATCTCAAGCT 180
QY 267 TATGTGATGCTCTTCTCATGCTTCACTTCACTGCTCTCAAGTT 314
DB 181 TATGTGATGCTCTTCTCATGCTTCACTTCACTGCTCTCAAGTT 228

RESULT 14
LOCUS AX411605 1302 bp DNA linear PAT 14-JUN-2002
DEFINITION Sequence 5 from Patent WO0229022.
ACCESSION AX411605
VERSION AX411605.1 GI:21444161
KEYWORDS
SOURCE Zea mays
ORGANISM Zea mays
Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
1 Thorpe, C., Cahoon, E.B. and Cahoon, R.E.
AUTHORS Chlorophyllase
TITLE Patent: WO 0229022-A 5 11-APR-2002;
JOURNAL E.I. DU PONT DE NEMOURS AND COMPANY (US)
FEATURES
LOCATION/Qualifiers
SOURCE 1..1302
/organism="Zea mays"
/mol_type="unassigned DNA"
/db_xref="taxon:4577"

ORIGIN
Query Match 17.8%; Score 201.6; DB 6; Length 1302;
Best Local Similarity 54.6%; Pred. No. 5.5e-39;
Matches 449; Conservative 0; Mismatches 364; Indels 9; Gaps 2;

QY 143 CTTCTAGAGCTTACCGCTTCCGCCAAGACAGCTGTGTGCTACCGCGGTGAGAGAG 202
DB 241 CTTCTAGAGCTTACCGCTTCCGCCAAGACAGCTGTGTGCTACCGCGGTGAGAGAG 300
QY 203 GAGATTATCCGGTGTGATGCTCTCATGCTTCACTTCTTACAACTCTTCTATCTTC 262
DB 301 GGGAGTACCGGATGCTGTCTCTACACGGCTACCTGCGGTGAACTCTTCACTCC 360
QY 263 AGCTATGTTGATGCTCTTCTCATGCTTCACTTCACTGCTCTCAAGTTATATAGTA 322
DB 361 AGCTATGTTGATGCTCTTCTCATGCTTCACTTCACTGCTCTCAAGTTATATAGTA 420
QY 323 TCGCCGACGACGACATGATGATTAATCAAGCGGAGATTATGATTTGTTAT 382
```

```
DB 421 TATCTGGGCGCGACACCGAGAGATCACTAGGGGGCGGTATGACTGAGCTAG 480
QY 383 CAGTAGAGACTTAAATCACTTCTTCCACCGCAAGTAAACCAAACTTACAAATTTGCC 442
DB 481 CCACGGGGCTGCGCTCAACTCTGCACTCGCGGTGCGGAACTTAAACCAAGTTGCCA 540
QY 443 TCTCGGCGCATGACCGCGGTGCGCAAAACCGGTTTGGGTGCGCTTAAAGAAATTTGGT 502
DB 541 TCTCGGCGCACAGTGGCGGCGGAGAGGTGGCTGTGCGGTGGCGCAACCCCAAG 600
QY 503 ACTCTCGAATCTTAAATCTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCA 562
DB 601 CCAAGCTGCTGCTCTCTCTGCGCGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCG 660
QY 563 AAGGAAACCAAAACCCCTCTCCGGTGTGCTTCACTTCACTTCACTTCACTTCACTTCA 622
DB 661 TGGGCAAGACAGACACCGCGCGCATCTTCACTTCACTTCACTTCACTTCACTTCACT 717
QY 623 AAACGCTTATCTTGTGATGCTTGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCG 682
DB 718 GTGCCCCCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 777
QY 683 CACCGTGTGACCTCCCGGAGTGAATGACGAGATTTCT-----TTGGGAATGTCAAG 736
DB 778 CGCGTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 837
QY 737 GTCCAGATGAGCATTTTGTGCGAAGATTAATGAGCATTTTGAAGATGATGATGATGAT 796
DB 838 CGCAGAGTGTCCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 897
QY 797 CAAAAGGATTAAGAGGAGAGATTTTATTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 856
DB 898 CGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 957
QY 857 CAATGAGGAGATTTGTTGTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 916
DB 958 CAAAGCGCGCGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1017
QY 917 ATGATCGTGAATTAGTTAAGATCAAGATGAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 958
DB 1018 ACGCGCGCGCGAGTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1059

RESULT 15
LOCUS AF337544 1115 bp mRNA linear PLN 22-OCT-2002
DEFINITION Brassica oleracea chlorophyllase 1 mRNA, complete cds.
ACCESSION AF337544
VERSION AF337544.1 GI:24210532
KEYWORDS
SOURCE Brassica oleracea
ORGANISM Brassica oleracea
Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE
1 (bases 1 to 1115)
AUTHORS Chen, H.-H., Chou, Y.-F. and Shaw, J.-F.
TITLE Molecular characterization and differential regulation of three
JOURNAL chlorophyllase genes, members of the novel serine esterase gene
family, in broccoli
Unpublished
2 (bases 1 to 1115)
AUTHORS Chen, H.-H., Chou, Y.-F. and Shaw, J.-F.
TITLE Direct Submission
JOURNAL Submitted (18-JAN-2001) Institute of Botany, Academia Sinica,
Taipei, Taiwan 11529, Republic of China
FEATURES
LOCATION/Qualifiers
SOURCE 1..1115
/organism="Brassica oleracea"
/mol_type="mRNA"
/cultivar="Green King"
/db_xref="taxon:3712"
```

5'UTR
CDS
/feature_type="f1orec"
1..6
7..981
/EC_number="3.1.1.14"
/note="BOC1H1"
/codon_start=1
/product="chlorophyllase 1"
/protein_id="AAN51933.1"
/db_xref="GI:24210533"
/translation="MAGKDESETFPSAATPLAFELGSLPTTVIPADPSANDLTAPKP
VILTSPTVAGTPVPLFPHGFVLRNFYSDIVINVAISHGYIVVAPOLCKILPBGQVE
VDDAGKVNMTSKNLKALPSSVANGNVTAIVGSHRGKTAFAVALGHAATLDSIK
FSALVGIDPVAGISKILITDEPEILTYKPEFDLMPVAVIGTGPKSNMLMPCAPA
EVNHEEFLIECKATGHRVADYGHMDLNDNLPGVGFPMAGCMCKNGKRSRSE
VGGIVVAPLKYSINGENSEIRQILKDPSEVSPARLDPEPEEASGYLV"
982..1115

3'UTR
ORIGIN

Query Match 16.2%; Score 183.4; DB 15; Length 1115;
Best Local Similarity 55.7%; Pred No. 1.9e-34;
Matches 437; Conservative 0; Mismatches 336; Indels 12; Gaps 4;

Oy 139 CCGTCTTAGAGCTTCAACCGTCTCCGCAAGACAGCTGTGGCTACGCCGGTGGAG 198
Db 103 CCGTGGCAACCGATTGTACCGCACTCCAAAGCCTGTAATATCATCCTCCCAACGCTC 162
Oy 199 GAAGAGATTATCCGGTGTGATGCTCCTCCATGGTTACCTTCTTACACTCCTTCTAT 258
Db 163 GCCGGAACCTTACCCCGTGTCTTATCTTCATGATTCATCTTCGTAACCTTCTAC 222
Oy 259 TCTGACTTATGTTGATGCTCTCTCTCATGGCTTCATCTCATGCTCTCATGTTATAT 318
Db 223 TCTGATGTTATTAACAGTAGCTTCTCATGGCTACATGTTGTAGCCCAAGCTTTC 282
Oy 319 AGTATGCCGGAACGACACAATGGATGATTAATCAACGGCGAGATTATGATGG 378
Db 283 AAGATTTCGCGCGGAGGGCAAGTGAAGTGAAGTCTGMAAAGTGTAAACTGG 342
Oy 379 TTATCAGTAGACTTAATCATCTTCTTCAGGGCAAGTAACCAACCTATCCAAATT 438
Db 343 ACTTCGAAAACCTCAAGCTCACTCCCAAGTTCAATCAAGTAATGGCAACTACAC 402
Oy 439 GCCCTCTCCGGCATAGCCGGGTGGCAAAACCGGTTGGCGTCCCTTAAAGAAATT 498
Db 403 GCACCTCGGGTCAATGCGCGGTGTAAACCGGTTGGCGTTCGTTAGGCCACGCC 462
Oy 499 G--GGTACTCTCGAATCTAAAGTCTGAATGATCGGTATGATCAAGTCAGTGA 555
Db 463 GCAACACTAGACCCATCCATCAAGTTTCAGCTTTGAGGAATGATCCAGTTGCAAGA 522
Oy 556 ACAGGGAAGGGAACCAACCCCTCTCCGGTGTGGCTTACCTTCCAACTCATTTGAC 615
Db 523 ATCAGCAATATGATAGAACCGATCCGAAATCTTAACGTACAAACCGGAATCATTCGAC 582
Oy 616 CTAGCAAAACGCCCTATCTTGTGATCGGTTGCGGGCTTGGTGAACCGCTCGAACCA 675
Db 583 CTGGACATGCCGGTTGCA---CTGATCGGTACGGGCTCGG--ACCGAAGATACATG 636
Oy 676 TTATTTCCACCGGTGCACTTCCCGAGTGAATCACCGAAGTTCTTTCGGGAATGCA 735
Db 637 CTGATGCCACATGCGCACACAGCGGAAGTGAACCATGAGAGTTTATATGAGTGAAG 696
Oy 736 GGTCCAGATGGCATTTTGGTGGCAAGATTATGCGCATTTGGACATGCTTGATGATGAT 795
Db 697 GCTAGGAAGGACATTTGCTGCTGCGGATTACGGAATATGATGATATGTGGACATAT 756
Oy 796 ACAAAAGGATTTAGAGGAAGATTCTTATGTTGTGTAAGATGTGAA--GAGAGG 852
Db 757 TTGCCCGGTTTGTGGGTTTATGCGCGGTTCTATGTGAACGCTAAACGCAAAAG 816
Oy 853 AGACCAATGAGAGATTGTTGGTGGACTTGTGTATCATTTTGAAGGCTTATTTGAA 912
Db 817 AGTGAGATGAGAACTTGTGTGGTGAATGTGGTTCGTTCTAAAGTATATGATATGG 876

Oy 913 GGAGA 917
Db 877 GGTA 881

Search completed: March 20, 2006, 16:11:53
Job time : 4016.27 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 20, 2006, 13:57:18 ; Search time 5143.9 Seconds

(Without alignments)
10323.556 Million cell updates/sec

Title: US-10-634-548-19

Perfect score: 1135
Sequence: 1 aaaaaaacgtcaagaaagaa.....aaaaaaaaaaaaaaaa 1135

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: gb_est1.*
2: gb_est2.*
3: gb_est3.*
4: gb_est4.*
5: gb_est5.*
6: gb_est6.*
7: gb_est7.*
8: gb_est8.*
9: gb_est9.*
10: gb_est10.*
11: gb_est11.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	622.2	54.8	652	1	AU237094 AU237094
2	539	47.5	641	9	AF005802 AF005802
3	525	46.3	624	9	AF106729 AF106729
4	507.8	44.7	883	10	CI497740 CI497740
5	399	35.2	399	3	BP788041 BP788041
6	398	35.1	413	1	AU228115 AU228115
7	375	33.0	427	3	BP786753 BP786753
8	347.6	30.6	481	6	CA992230 CA992230
9	341.6	30.1	494	6	CA992233 CA992233
10	325	28.6	749	8	CK076457 CK076457
11	312.4	27.5	504	11	CR397401 CR397401
12	308.6	27.2	317	1	AV356714 AV356714
13	299.2	26.4	542	7	CN736070 CN736070
14	298	26.3	827	8	CK076456 CK076456
15	287	25.3	842	8	DR932000 DR932000
16	283.4	25.0	761	6	CA815004 CA815004
17	261.2	23.0	492	11	CR397400 CR397400
18	259	22.8	613	1	AJ808610 AJ808610
19	258.8	22.8	790	6	CF440994 CF440994
20	253.6	22.3	569	6	CV525268 CV525268
21	236.4	20.8	811	7	CO985817 CO985817
22	224	19.7	611	7	CN490540 CN490540

23	220.2	19.4	510	7	CV525641 CV525641
24	213	18.8	776	6	CA816983 CA816983
25	206.2	18.2	848	8	DR533022 DR533022
26	205.6	18.1	858	10	CZ950599 CZ950599
27	205.4	18.1	634	3	BP955021 BP955021
28	198.6	17.5	818	8	DR449908 DR449908
29	195.8	17.3	551	1	AW039135 AW039135
30	195	17.2	724	3	BQ113489 BQ113489
31	193.6	17.1	499	3	BI699790 BI699790
32	191	16.8	588	6	CD826708 CD826708
33	189.4	16.7	577	3	BQ113488 BQ113488
34	184	16.2	649	8	DR095974 DR095974
35	183.8	16.2	865	8	DR015096 DR015096
36	182	16.0	758	9	BH511470 BH511470
37	178.6	15.7	628	3	BP957354 BP957354
38	175.2	15.4	750	8	DR539554 DR539554
39	172.2	15.2	497	6	CF608740 CF608740
40	169	14.9	541	3	BM885893 BM885893
41	164.8	14.5	668	7	CV187182 CV187182
42	162.6	14.3	553	3	BM885730 BM885730
43	154.8	13.6	455	2	BE923513 BE923513
44	154.6	13.6	715	8	DR534907 DR534907
45	149.6	13.2	699	7	CO487862 CO487862

ALIGNMENTS

RESULT 1
AU237094 652 bp mRNA linear EST 01-APR-2002
DEFINITION AU237094 RAPL15 Arabidopsis thaliana cDNA clone RAPL15-48-F08 5', mRNA sequence.
ACCESSION AU237094
VERSION AU237094.1 GI:19876263
KEYWORDS EST.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE
Akiyama, K., Enju, A., Oono, Y., Sakurai, T., Carninci, P., Kawai, J., Itoh, M., Ishii, Y., Arikawa, T., Shibata, K., Shinagawa, A., Muramatsu, M., Hayashizaki, Y. and Shinozaki, K.
Large scale analysis of Arabidopsis full-length cDNA
Unpublished (2002)
Contact: Motoaki Seki
Plant Functional Genomics Research Group
RIKEN Genomic Sciences Center
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-4359
Fax: 81-298-36-9060
Email: msek@rc.riken.go.jp
An Arabidopsis full-length cDNA library was constructed essentially as reported previously (Seki et al., 1998). cDNA cleaved with BamHI and XhoI was ligated to modified lambda FIC-1 vector (Carninci et al., submitted for publication) digested with BamHI and SalI. This clone is in a modified pluscript vector. Please visit our web site (http://www.gsc.riken.go.jp/e/plant/index_e.html) for further details.

TITLE
JOURNAL
COMMENT
FEATURES
source
Location/Qualifiers
1..652
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="RAPL15-48-F08"
/ciseue_type="mixture of silique and flower"
/lab_host="DH10B"
/clone_lib="RAPL15"
/note="Site_1: BamHI; Site_2: SalI"

ORIGIN

Query Match 54.8%; Score 622.2; DB 1; Length 652;
 Best Local Similarity 99.1%; Pred. No. 4e-146; Matches 646; Conservative 0; Mismatches 4; Indels 2; Gaps 2;

```

Oy 117 ATCTGCTGTCGCAAAATTAACACCGCTCTTCTAGAGCTTACCGCTCCGCCAAGACG-C 175
Db 1 ATCTGCTGTCGCAAAATTAACACCGCTCTTCTAGAGCTTACCGCTCCGCCAAGACG-C 60
Oy 176 TGTGCTGCTGTCGCAAAATTAACACCGCTCTTCTAGAGCTTACCGCTCCGCCAAGACG 235
Db 61 TGTGCTGCTGTCGCAAAATTAACACCGCTCTTCTAGAGCTTACCGCTCCGCCAAGACG 120
Oy 236 ACCCTCTGTCGCAAAATTAACACCGCTCTTCTAGAGCTTACCGCTCCGCCAAGACG 295
Db 121 ACCCTCTGTCGCAAAATTAACACCGCTCTTCTAGAGCTTACCGCTCCGCCAAGACG 180
Oy 236 TCCCTCATGCTCTCTGATTAATAGTATGCGCGACGACGACGACGATGATTAAT 355
Db 181 TCCCTCATGCTCTCTGATTAATAGTATGCGCGACGACGACGACGATGATTAAT 240
Oy 356 CAACGCGGAGATTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 415
Db 241 CAACGCGGAGATTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
Oy 416 TAACGCAAAACCTATCCAAATTTGCGCTCTCGCGCATAGCGCGGCGCAAAACCGGCT 475
Db 301 TAACGCAAAACCTATCCAAATTTGCGCTCTCGCGCATAGCGCGGCGCAAAACCGGCT 360
Oy 476 TTGCGGTGCGCTTAAAGAAATTTGGGTACTCTCGATCTTAAAGATCTCGACATTTGATCG 535
Db 361 TTGCGGTGCGCTTAAAGAAATTTGGGTACTCTCGATCTTAAAGATCTCGACATTTGATCG 420
Oy 536 GTATGATTCAGTCGATGGAACAGGGAAGGGAACAAACCCCTCTCGGCTTGGGCTT 595
Db 421 GTATGATTCAGTCGATGGAACAGGGAAGGGAACAAACCCCTCTCGGCTTGGGCTT 480
Oy 596 ACCCTCAAACTCATTTGACCTAGACAAACGCTATCTTGTGATCGGCTTGGGCTT 655
Db 481 ACCCTCAAACTCATTTGACCTAGACAAACGCTATCTTGTGATCGGCTTGGGCTT 540
Oy 656 GTGAAACGCTGGAACCCATTTATCCACCGGTGACCT-CCGAGATGATCACCGA 714
Db 541 GTGAAACGCTGGAACCCATTTATCCACCGGTGACCT-CCGAGATGATCACCGA 600
Oy 715 GAGTCTTTTGGGATGTCAGATGTCAGATGTCAGATGTCAGATGTCAGATGTCAGAT 766
Db 601 GAGTCTTTTGGGATGTCAGATGTCAGATGTCAGATGTCAGATGTCAGATGTCAGAT 652

```

RESULT 2
 AF005802 641 bp DNA linear GSS 06-NOV-2000
 LOCUS AF005802 Arabidopsis thaliana 188-4 Arabidopsis thaliana genomic
 DEFINITION clone 1884xa similar to EST with Genbank Accession Number R30366,
 genome survey sequence.
 ACCESSION AF005802
 VERSION AF005802.1 GI:3387726
 KEYWORDS GSS;
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 REFERENCE 1 (Baes 1 to 641)
 Machur, J., Szabados, L., Schaefer, S., Gruenberg, B., Lossow, A.,
 Jonas-Straube, B., Scheil, J., Koncz, C., and Koncz-Kalman, Z.
 TITLE Gene identification with sequenced T-DNA tags generated by
 JOURNAL transformation of Arabidopsis cell suspension
 PUBMED Plant J. 13 (5), 707-716 (1998)
 COMMENT 9681013
 Contact: Koncz C
 Abteilung Genetische Grundlagen der Pflanzenzuechtung

Max-Planck Institut fuer Zuechtungsforchung
 Carl von Linné weg 10, Cologne, D-50829, Germany
 Email: koncz@mpiz-koeln.mpg.de
 transgenic cell line was obtained by transformation with the T-DNA
 of pPCV6NFHY Agrobacterium binary vector; the left border junction
 of T-DNA insertion 1884xa was isolated in E. coli after XbaI
 digestion and self-circularization of plant DNA; clone 1884xa
 carries a plant DNA fragment of 5.0 kb that extends from a XbaI
 site to the left-border junction of pPCV6NFHY T-DNA tag; sequences
 of the left T-DNA border are excluded from the submission
 Class: transposon-tagged.
 Location/Qualifiers
 1..641
 /organism="Arabidopsis thaliana"
 /mol_type="genomic DNA"
 /ecotype="Col-1"
 /db_xref="taxon:3702"
 /clone="1884xa"
 /cell_line="188-4"
 /clone_1fb="Arabidopsis thaliana 188-4"

ORIGIN

Query Match 47.5%; Score 539; DB 9; Length 641;
 Best Local Similarity 99.5%; Pred. No. 4.3e-125; Matches 550; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

```

Oy 310 CAGTTATATGATTCGCGGACCGACACAAATGATGATGATTAATCAACGCGGAGATT 369
Db 90 CAGTTATATGATTCGCGGACCGACACAAATGATGATGATTAATCAACGCGGAGATT 149
Oy 370 ATGATTTGTTATCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 429
Db 150 ATGATTTGTTATCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 209
Oy 430 TCCAATTTGCTCTCTCGGACCATAGCGCGGTGCAAAACCGGTTGCGCTTGA 489
Db 210 TCCAATTTGCTCTCTCGGACCATAGCGCGGTGCAAAACCGGTTGCGCTTGA 269
Oy 490 AAGAAATTTGGGTACTCTCGAATCTTAAAGATCTCGACATTTGATGATGATGATGAT 549
Db 270 AAGAAATTTGGGTACTCTCGAATCTTAAAGATCTCGACATTTGATGATGATGATGAT 329
Oy 550 GATGGAACAGGGAAGGGAACAAACCCCTCTCGGCTTGGCTTACCTTCAAACTCA 609
Db 330 GATGGAACAGGGAAGGGAACAAACCCCTCTCGGCTTGGCTTACCTTCAAACTCA 389
Oy 610 TTGACCTGACAAACGCTTATCTTGTATGCTTGGGCTTGGTGAACCGCTCG 669
Db 390 TTGACCTGACAAACGCTTATCTTGTATGCTTGGGCTTGGTGAACCGCTCG 449
Oy 670 AACCATTTATCCACCGGTGACCTCCGAGTGAATACCGAGAGTTCTTCCGGA 729
Db 450 AACCATTTATCCACCGGTGACCTCCGAGTGAATACCGAGAGTTCTTCCGGA 509
Oy 730 TGTCAAGTCCAGCATGCTTTCGTTGCAAGGATTTAGGCACTTTGACATCTTGAT 789
Db 510 TGTCAAGTCCAGCATGCTTTCGTTGCAAGGATTTAGGCACTTTGACATCTTGAT 569
Oy 790 GATGATCAAAAGGATTTAAGGAGAGTTCTTATTTGTTGTGAATGATGGAAG 849
Db 570 GATGATAC-NNAGGATTTAAGGAGAGTTCTTATTTGTTGTGAATGATGGAAG 628
Oy 850 AGGAGACCAATGA 862
Db 629 AGGAGACCAATGA 641

```

RESULT 3
 AF106729 624 bp DNA linear GSS 06-NOV-2000
 LOCUS AF106729 Arabidopsis thaliana Col-0 Arabidopsis thaliana genomic
 DEFINITION clone 031-2XA.PC3, genomic survey sequence.
 ACCESSION AF106729

VERSION AF106729.1 GI:4588400
 KEYWORDS GSS.
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 rosids; eurosid II; Brassicales; Brassicaceae; Arabidopsids.
 1 (bases 1 to 624)
 REFERENCE Koncz, C., Szabados, L., Grunberg, B. and Schaefer, S.
 TITLE Gene identification with sequenced T-DNA tags in Arabidopsis
 JOURNAL Unpublished (1998)
 COMMENT Contact: Koncz C
 Abteilung Genetische Grundlagen der Pflanzenzüchtung
 Max-Planck Institut fuer Zuechtungsforsehung
 Carl von Linné weg 10, Cologne, D-50829, Germany
 Email: koncz@mpiz-koeln.mpg.de
 T-DNA tagged line 031-2, PCR fragment 031-2XA; Transgenic plant
 line 031-2 was obtained by transformation with the T-DNA of
 pPCV6NFHYg Agrobacterium binary vector; the left border junction of
 T-DNA insertion 031-2XA was isolated by PCR amplification using the
 LB2 (5'-CTGGGAATGGCGAAATCAAGGATGATGTAAG-3') and PC3
 (5'-CCTTGGCCCTGAGTGGCTTGGCGACG-3') primers and sequenced with the
 PC3 primer as described by Mathur et al. Plant J. (1998) 13.
 707-716; the PCR-amplified DNA fragment extends from an XbaI site
 to the left border junction of pPCV6NFHYg T-DNA tag, sequences of
 the left T-DNA end are excluded from the submission
 Class: transposon-tagged.
 FEATURES
 source
 1..624
 /organism="Arabidopsis thaliana"
 /mol_type="genomic DNA"
 /ecotype="Col-0"
 /db_xref="taxon:3702"
 /clone="031-2XA.PC3"
 /clone_1lb="Arabidopsis thaliana Col-0"
 ORIGIN
 Query Match 46.3%; Score 525; DB 9; Length 624;
 Best Local Similarity 100.0%; Pred. No. 1,5e-121;
 Matches 525; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 790 GATGATCAAAAGGATTAAGGAGAGTTCTTATTGTTGCT 834
 DB 580 GATGATCAAAAGGATTAAGGAGAGTTCTTATTGTTGCT 624
 RESULT 4
 LOCUS CL497740/c 883 bp DNA linear GSS 01-APR-2004
 DEFINITION SAIL_646_E09.v1 SAIL Collection Arabidopsis thaliana genomic clone
 SAIL_646_E09.v1, genomic survey sequence.
 ACCESSION CL497740
 VERSION CL497740.1 GI:45992433
 KEYWORDS GSS.
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 rosids; eurosid II; Brassicales; Brassicaceae; Arabidopsids.
 1 (bases 1 to 883)
 REFERENCE Sessions, A., Burke, E., Presting, G., Aux, G., McElver, J., Patton, D.,
 Dietrich, B., Ho, P., Bacwaden, J., Ko, C., Clarke, J. D., Cotton, D.,
 Bullis, D., Snell, J., Miquel, T., Hutcheson, D., Kimmerly, B.,
 Mitzel, T., Katsirji, F., Glazebrook, J., Law, M. and Goff, S.A.
 TITLE A high-throughput Arabidopsis reverse genetics system
 JOURNAL Plant Cell 14 (12), 2985-2994 (2002)
 PUBMED 12468722
 COMMENT Contact: Sessions A
 Applied Trait Genetics
 Syngenta Biotechnology Inc.
 3054 Cornwallis Rd., Research Triangle Park, NC 27709, USA
 Email: allen.sessions@syngenta.com
 ARRC Stock Number CS827897; T-DNA left border flanking sequences of
 Syngenta Arabidopsis insertion library (SAIL) lines are available
 through the Arabidopsis Biological Resource Center (ABRC).
 Sequences represent a pool of amplified genomic regions and not
 single contiguous sequences.
 Class: T-DNA tagged.
 FEATURES
 source
 1..883
 /organism="Arabidopsis thaliana"
 /mol_type="genomic DNA"
 /ecotype="Columbia"
 /db_xref="taxon:3702"
 /clone="SAIL_646_E09.v1"
 /clone_1lb="SAIL Collection"
 /note="T-DNA left border sequences were isolated using a
 modified TAIL-PCR strategy"
 ORIGIN
 Query Match 44.7%; Score 507.8; DB 10; Length 883;
 Best Local Similarity 96.8%; Pred. No. 3.6e-117;
 Matches 518; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 651 GCTTGTAAGAAACCGCTGGAAACCATTAATCCACGCTGTCACCTCCCGAGTGAATCA 710
 DB 371 GCTTGTAAGAAACCGCTGGAAACCATTAATCCACGCTGTCACCTCCCGAGTGAATCC 312
 QY 711 CCGAGAGTTCTTTCGGGAATGTCAGAGTCCAGCATGAGATTTCGTTGGGAAGATTATAG 770
 DB 311 CCGAGAGTTCTTTCGGGAATGTCAGAGTCCAGCATGAGATTTCGTTGGGAAGATTATAG 252
 QY 771 GCATTGGACATGCTTGAATGATGATACAAAGGATTGAGGAAAGATTCTTATGTTT 830
 DB 251 GCATTGGACATGCTTGAATGATGATACAAAGGATTGAGGAAAGATTCTTATGTTT 192
 QY 831 GTGTAAAGATGAGTGAAGAGAGAGACCAATGAGAGATTGCTGTTGGTGAATTGTT 885
 DB 191 GTGTAAAGATGAGTGAAGAGAGAGACCAATGAGAGATTGCTGTTGGTGAATTATTT 137

RESULT 5
 BP789041/c 399 bp mRNA linear EST 10-FEB-2005
 LOCUS BP789041 RAFL7 Arabidopsis thaliana cDNA clone RAFL07-36-K05 3',
 DEFINITION mRNA sequence.
 ACCESSION BP789041 GI:59268282
 VERSION BP789041
 KEYWORDS EST.
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana (thale cress)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 1 (bases 1 to 399)
 REFERENCE Seki, M., Ishida, J., Kamiya, A., Satou, M., Nakajima, M., Akiyama, K.,
 Iida, K., Enju, A., Sakurai, T., Arakawa, T., Carninci, P., Fukuda, S.,
 Iida, J., Kawai, J., Sasaki, D., Shiraki, T., Hayashizaki, Y. and
 Shinozaki, K.
 Large-Scale Analysis of RIKEN Arabidopsis Full-length cDNAs
 Unpublished (2005)
 TITLE Contact: Motoaki Seki
 JOURNAL Plant Functional Genomics Research Group
 COMMENT RIKEN Genomic Sciences Center
 3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
 Tel: 81-298-36-4359
 Fax: 81-298-36-9060
 Email: mseki@rtc.riken.go.jp
 An Arabidopsis full-length cDNA library was constructed essentially
 as reported previously (Seki et al., 1998, 2002). This clone is in a
 modified Bluescript vector.
 Please visit our web site (http://pfweb.gsc.riken.jp and
 http://large.gsc.riken.jp) for further details.
 reversed clone.

FEATURES
 source location/Qualifiers
 1..399
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /db_xref="taxon:3702"
 /clone="RAFL07-36-K05"
 /dev_stage="rosette plants"
 /lab_host="DH10B"
 /clone_1lb="RAFL7"
 /note="Site 1: BamHI; Site 2: SalI; subjected to
 cold-treated (1, 2, 5, 10, 24 hr)"

ORIGIN
 Query Match 35.2%; Score 399; DB 3; Length 399;
 Best Local Similarity 100.0%; Pred. No. 1e-89;
 Matches 399; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 673 CCATATATCCACGCTGTCACCTCCCGAGTGAATCCAGAGTTCTTTCGGGAATGT 732
 DB 399 CCATATATCCACGCTGTCACCTCCCGAGTGAATCCAGAGTTCTTTCGGGAATGT 340
 QY 733 CAAGGTCCAGCATGAGTTCGTTGCGAAGATTATGAGCATTTGACATGCTTGATAT 792

DB 339 CAAGGTCCAGCATGAGTTCGTTGCGAAGATTATGAGCATTTGACATGCTTGATAT 280
 QY 793 GATTCAAAAGGATTTAAGGAAAGATTTTATTTGTTGTAAGATTGTAAGAGAG 852
 DB 279 GATTCAAAAGGATTTAAGGAAAGATTTTATTTGTTGTAAGATTGTAAGAGAG 220
 QY 853 AGACCAATGAGAGATTCGTTGGGACCTGTTGATCATTTTGAAGGCTTATTTGAA 912
 DB 219 AGACCAATGAGAGATTCGTTGGGACCTGTTGATCATTTTGAAGGCTTATTTGAA 160
 QY 913 GGAATGATCGTGAATTAGTTAAGATCAAGATGGGTGTCAGAGATGTTCCGTTGAA 972
 DB 159 GGAATGATCGTGAATTAGTTAAGATCAAGATGGGTGTCAGAGATGTTCCGTTGAA 100
 QY 973 ATTCAAGATTGAGTTATCATGTAACATTAAGTTTCTTTAGGGGCTGTTTCTA 1032
 DB 99 ATTCAAGATTGAGTTATCATGTAACATTAAGTTTCTTTAGGGGCTGTTTCTA 40

RESULT 6
 AU228115/c 413 bp mRNA linear EST 23-APR-2002
 LOCUS AU228115 RAFL15 Arabidopsis thaliana cDNA clone RAFL15-48-F08 3',
 DEFINITION mRNA sequence.
 ACCESSION AU228115
 VERSION AU228115
 KEYWORDS AU228115.1 GI:19742762
 SOURCE EST.
 ORGANISM Arabidopsis thaliana (thale cress)
 Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 1 (bases 1 to 413)
 REFERENCE Seki, M., Narusaka, M., Ishida, J., Kamiya, A., Satou, M., Nakajima, M.,
 Akiyama, K., Enju, A., Oono, Y., Sakurai, T., Carninci, P., Kawai, J.,
 Itoh, M., Ishii, Y., Arakawa, T., Shibata, K., Shinozaki, K.,
 Muramatsu, M., Hayashizaki, Y. and Shinozaki, K.
 Large scale analysis of Arabidopsis full-length cDNA
 Unpublished (2002)
 TITLE Contact: Motoaki Seki
 JOURNAL Plant Functional Genomics Research Group
 COMMENT RIKEN Genomic Sciences Center
 3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
 Tel: 81-298-36-4359
 Fax: 81-298-36-9060
 Email: mseki@rtc.riken.go.jp
 An Arabidopsis full-length cDNA library was constructed essentially
 as reported previously (Seki et al., 1998). cDNA cleaved with BamHI
 and XhoI was ligated to modified lambda PUC-1 vector (Carninci et
 al., submitted for publication) digested with BamHI and SalI. This
 clone is in a modified Bluescript vector. Please visit our web
 site (http://www.gsc.riken.go.jp/e/plant/index_e.html) for further
 details.

FEATURES
 source location/Qualifiers
 1..413
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /db_xref="taxon:3702"
 /clone="RAFL15-48-F08"
 /tissue_type="mixture of silique and flower"
 /lab_host="DH10B"
 /clone_1lb="RAFL15"
 /note="Site 1: BamHI; Site 2: SalI"

ORIGIN
 Query Match 35.1%; Score 398; DB 1; Length 413;
 Best Local Similarity 98.8%; Pred. No. 1.8e-89;
 Matches 401; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 715 GAGTCTCTCGGAGATGTCAAGTCCAGCATGGCATTTGTTGGAGAGATTAATGGGAT 774
DB 413 GAGTCTCTCGGAGATGTCAAGTCCAGCATGGCATTTGTTGGAGAGATTAATGGGAT 354
QY 775 TTGGACATGCTTGTATGATATACAAAGGATTAAGAGGAGAGTCTTATTTGTTGT 834
DB 353 TTGGACATGCTTGTATGATATACAAAGGATTAAGAGGAGAGTCTTATTTGTTGT 294
QY 835 AAGATGTTGAAGAGAGAGACCAATGAGGAGATTCGTTGTGATCTTTGTATCATTT 894
DB 293 AAGATGTTGAAGAGAGAGACCAATGAGGAGATTCGTTGTGATCTTTGTATCATTT 234
QY 895 TTGAAGCTTATTTGAAGAGAGATGATGTAATTAAGTAAAGATGGATGATC 954
DB 233 TTGAAGCTTATTTGAAGAGAGATGATGTAATTAAGTAAAGATGGATGATC 174
QY 955 GAGATGTTCCGTTGAATTAATCAAGATTTGAGTTATCATGTAACATTAAGTTTCTT 1014
DB 173 GAGATGTTCCGTTGAATTAATCAAGATTTGAGTTATCATGTAACATTAAGTTTCTT 114
QY 1015 TAGGGGCTGTTTCTTATTTGATCATATCATCATGCTTTGTTGCTTATGTTTACAAC 1074
DB 113 TAGGGGCTGTTTCTTATTTGATCATATCATCATGCTTTGTTGCTTATGTTTACAAC 54
QY 1075 TTATATTGTAACACTTTTAAGTACACCTCTTGTCTTAACAAAAA 1120
DB 53 TTATATTGTAACACTTTTAAGTACACCTCTTGTCTTATGATATTAA 8

RESULT 7
BP786753/c 427 bp mRNA linear EST 10-FEB-2005
LOCUS BP786753 Arabidopsis thaliana cDNA clone RAFL26-01-124 3,
DEFINITION mRNA sequence.
ACCESSION BP786753
VERSION BP786753.1 GI:59263657
KEYWORDS EST.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana

REFERENCE
AUTHORS Seki,M., Ishida,J., Kamiya,A., Setou,M., Nakajima,M., Akiyama,K.,
Iida,K., Enju,A., Sakurai,T., Arakawa,T., Carninci,P., Fukuda,S.,
Iida,J., Kawai,J., Sasaki,D., Shiraki,T., Hayashizaki,Y. and
Shinozaki,K.

TITLE Large-Scale Analysis of RIKEN Arabidopsis Full-length cDNAs
JOURNAL Unpublished (2005)
COMMENT Contact: Motoaki Seki
Plant Functional Genomics Research Group
RIKEN Genomic Sciences Center
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-4359
Fax: 81-298-36-9060
Email: msek@rcc.riken.go.jp

FEATURES
source
1. 427
Location/Qualifiers
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="RAFL26-01-124"
/dev_stage="rosea plants"
/lab_host="DH10B"
/clone_id="RAFL7"
/note="Site 1: BamHI, Site 2: SalI, subjected to
cold-treated (1, 2, 5, 10, 24 hr)"

ORIGIN
Query Match 33.0%; Score 375; DB 3; Length 427;
Best Local Similarity 98.7%; Pred. No. 1,2e-83;
Matches 378; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 738 TCCAGCATGTCATTTCTTCCGAAGATTAATGGGATTTGACATGTTGATGATGATAC 797
DB 427 TCCAGCATGTCATTTCTTCCGAAGATTAATGGGATTTGACATGTTGATGATGATAC 368
QY 798 AAAAGGATTAAGAGAGAGAGTCTTATTTGTTGTGAATGTGTGAAGAGAGACC 857
DB 367 AAAAGGATTAAGAGAGAGAGTCTTATTTGTTGTGAATGTGTGAAGAGAGACC 308
QY 858 AATGAGAGATTCGTTGTGAGACTTGTGTATCATTTTGAAGCTTATTTGAAGAGA 917
DB 307 AATGAGAGATTCGTTGTGAGACTTGTGTATCATTTTGAAGCTTATTTGAAGAGA 248
QY 918 TGATCGGAATTTGTTAATGATCAAAAGTGGGTGACAGAGATGTTCCGTTGAATTC 977
DB 247 TGATCGGAATTTGTTAATGATCAAAAGTGGGTGACAGAGATGTTCCGTTGAATTC 188
QY 978 AGAGTTGAGTTATCATGTAACATTAATGTTTCTTGAAGGCTGGTTTCTATGTC 1037
DB 187 AGAGTTGAGTTATCATGTAACATTAATGTTTCTTGAAGGCTGGTTTCTATGTC 128
QY 1038 AATATCATGCTTTGTTGCTTATGTTTACAACTTATTTGTAACAATCTTTAAGT 1097
DB 127 AATATCATGCTTTGTTGCTTATGTTTACAACTTATTTGTAACAATCTTTAAGT 68
QY 1098 CACCTCTTGTCTTACAAAAA 1120
DB 67 CACCTCTTGTCTTATGATATTAA 45

RESULT 8
CA992230 481 bp mRNA linear EST 10-OCT-2003
LOCUS CA992230
DEFINITION HC0810 GIBCOBRL CAT. NO. 19643-014 Brassica rapa subsp. pekinensis
cDNA, mRNA sequence.
ACCESSION CA992230
VERSION CA992230.1 GI:37621525
KEYWORDS EST.
SOURCE Brassica rapa subsp. pekinensis
ORGANISM Brassica rapa subsp. pekinensis
REFERENCE
AUTHORS Gao,R.J., Dai,D.P. and Ma,R.C.
TITLE Expressed sequence tags of heading leaf during the heading process
JOURNAL Unpublished (2003)
COMMENT Contact: Rongcai Ma
Plant Functional Genomics
Beijing Agrobiotechnology Research Center
Haidian District, Ban-Jing Rd., Beijing 100089, China
Tel: 86 10 5150 3831
Fax: 86 10 5150 3980
Email: rcma1@yahoo.com
Seq primer: T7
High quality sequence stop: 481
POLYA=No.

FEATURES
source
1. 481
Location/Qualifiers
/organism="Brassica rapa subsp. pekinensis"
/mol_type="mRNA"
/sub_species="pekinensis"
/db_xref="taxon:5151"
/dev_stage="heading leaf"
/lab_host="E.coli DH10B(zip)"
/clone_id="GIBCOBRL CAT. NO. 19643-014"

/note="Vector: pZL1; Site 1: NotI; Site 2: SalI; cDNA library was constructed by SUPERScript™ Lambda System for cDNA Synthesis and Cloning (GIBCOBRL, CAT. NO.19643-014)"

ORIGIN

Query Match 30.6%; Score 347.6; DB 6; Length 481;
Best Local Similarity 84.3%; Pred. No.1e-76; Mismatches 74; Indels 1; Gaps 1;
Matches 403; Conservative 0; Mismatches 74; Indels 1; Gaps 1;

Oy 508 TCGAATCTTAAAGATCTCGACATTGATCGGTATGATTCAGTCGATGAAACAGGGAAAGG 567
Db 2 TCGGAATCTTAAAGATCTCGACATTGATCGGTATGATTCAGTCGATGAAACAGGGAAAGG 61
Oy 568 AAACAAACCCCTCCCGGCTGTTGGCTTACCTTCCAACTCATTTGACTGACAAACG 627
Db 62 AAACAAACCCCTCCCGGCTTAACTTAACTGAAACCAACTCATTTAACCTGAAAAAGATG 121
Oy 628 CCTATATCTTGATCGGTTCCGGGCTTGTAACCGCTCGAAACCCATTTATCCACCG 687
Db 122 CCTGTTCTAGTTATGTTGTTCCGGACTTGTGAACTTGCCTCGAAACCCATTTGTTCCACG 181
Oy 688 TGTGACCTCCCGAGTGAATCACCGAAGTTCTTTCCGAAATGTCAAGTCCAGCATGG 747
Db 182 TGTGACCTACGGGAGTGAACACCGAAGTTTTCAGAAATGTCAAGTCCAGCGTG 241
Oy 748 CATTTGTTGGGAAGATTAAGGAGATTGACATGCTTGAATGATATCAAAAGGAT 807
Db 242 CATTTGTTGGCAAGATTAAGGACATTGACATGCTTGAATGATATCAAAAGGAGCTT 301
Oy 808 AGAGGAGAGAGTTCTTATTTGTTGTGAAGATGTAAGAGAGAGAGACCAATGAGAGA 867
Db 302 AGAGGAGAGAGTTCTTATTTGTTGTGAAGATGTAAGAGAGAGAGAGAGAGAGAG 361
Oy 868 TTGCTGTGACCTGTTGATCATTTTGAAGCTTATTTGAAGAGAGATGATGTGA 927
Db 362 TTATTTGGGTATTTGTTGTTGCTTTTGAAGCTTATTTGAAGAGAGATGATGTGAG 421
Oy 928 TTAGTAAGATCAAAAGAT-GGGTGCACGAGAGATTCCTCCGTGAATTCAGAGATT 984
Db 422 TTGATGAAGATCAAAAGATGGGTGTGATGAAGTGTCTCTGTGAATTCAGAGATT 479

RESULT 9
CA992233 494 bp mRNA linear EST 10-OCT-2003
LOCUS HC0819 GIBCOBRL CAT. NO. 19643-014 Brassica rapa subsp. pekinensis
DEFINITION cDNA, mRNA sequence.
ACCESSION CA992233
VERSION CA992233.1 GI:37621528
KEYWORDS EST
SOURCE Brassica rapa subsp. pekinensis
ORGANISM Brassica rapa subsp. pekinensis
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
1 (bases 1 to 494)
REFERENCE
AUTHORS Gao,R.J., Dai,D.P. and Ma,R.C.
TITLE Expressed sequence tags of heading leaf during the heading process
JOURNAL Unpublished (2003)
COMMENT Contact: Rongcai, Ma
Beijing Agrobiotechnology Research Center
Haidian District, Ban-Jing Rd., Beijing 100089, China
Tel: 86 10 5150 3831
Fax: 86 10 5150 3980
Email: rcma1@yahoo.com
Seq primer: 17
High quality sequence stop: 494
POLY=A=No.

FEATURES
SOURCE 1..494
Location/Qualifiers
/organism="Brassica rapa subsp. pekinensis"

/mol_type="mRNA"
/sub_species="pekinensis"
/db_xref="taxon:51351"
/tissue_type="heading leaf"
/dev_stage="folding stage"
/lab_host="E.coli DH10B(ZIP)"
/clone_id="GIBCOBRL CAT. NO. 19643-014"
/note="Vector: pZL1; Site 1: NotI; Site 2: SalI; cDNA library was constructed by SUPERScript™ Lambda System for cDNA Synthesis and Cloning (GIBCOBRL, CAT.NO.19643-014)"

ORIGIN

Query Match 30.1%; Score 341.6; DB 6; Length 494;
Best Local Similarity 84.1%; Pred. No.3.4e-75; Mismatches 74; Indels 1; Gaps 1;
Matches 397; Conservative 0; Mismatches 74; Indels 1; Gaps 1;

Oy 508 TCGAATCTTAAAGATCTCGACATTGATCGG-TATGATCTCAGTCGATGAAACAGGGAAAGG 566
Db 2 TCGGAATCTTAAAGATCTCGACATTGATCGGTTGATGATCCGGTGAAGAGAGGAAAGG 61
Oy 567 GAAACAAACCCCTCCCGGCTTGGCTTACCTTCCAACTCATTTGACTGACAAAC 626
Db 62 TAAACAAACCCCTCCCGGCTTTTAACTTAACTGAAACCAACTTATTAACCTGAAAAAGAT 121
Oy 627 GCTATATCTTGATCGGTTCCGGGCTTGTAACCGCTCGAAACCCATTTATCCACG 686
Db 122 GCTGTTCTAGTTATGTTGTTCCGGACTTGTGAACTTGCCTCGAAACCCATTTGTTCCACG 181
Oy 687 GTGTGACCTCCCGAGTGAATCACCGAAGTTCTTTCCGGAATGTCAAGTCCAGCATG 746
Db 182 GTGTGACCTTACGGAGTGAACACCGAAGTTTTCAGGAATGTCAAGTCCAGCGTG 241
Oy 747 GCATTTGTTCCGAAGATTAAGGCAATTGAGATGCTTGAATGATATCAAAAGGAT 806
Db 242 GCATTTGTGCAAGAGATTAAGGCAATTGAGATGCTTGAATGATATCAAAAGGCTT 301
Oy 807 TAGAGGAGAGAGTTCTTATTTGTTGTGAAGATGTAAGAGAGAGAGACCAATGAGAG 866
Db 302 TAGAGGAGAGAGTTCTTATTTGTTGTGAAGATGTAAGAGAGAGAGAGAGAGAG 361
Oy 867 ATTGTTGTGACCTTGTGATCATTTTGAAGCTTATTTGAAGAGAGATGATGTGA 926
Db 362 ATTATTTGTTGATTTGTTGTTGCTTTTGAAGCTTATTTGAAGAGATGATGTGGA 421
Oy 927 ATTGTTTAAGTCAAAAGATGGGTGTACAGAGATGTTCCCTGAATTCAA 978
Db 422 GTTATGAAGATCAAAAGATGGGTGTGATGAAGAGTGTCTGTGTAATTCAA 473

RESULT 10
CX076457 749 bp mRNA linear EST 14-DEC-2004
LOCUS UCRCS08_50B04_g Parent Washington Navel Orange Callus cDNA Library
DEFINITION UCRCS08-2 Citrus sinensis cDNA clone UCRCS08-50B04-C8-1-5-9; mRNA
sequence.
ACCESSION CX076457
VERSION CX076457.1 GI:56590447
KEYWORDS EST
SOURCE Citrus sinensis
ORGANISM Citrus sinensis
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Sepidales; Rutaceae; Citrus.
1 (bases 1 to 749)
REFERENCE
AUTHORS Close,T.J., Roose,M.L., Ye,X.R., Fenton,R.D., Manamaker,S.,
Lyon,M., Jang,C., Quintilio,C., Ikeda,J., Collin,M., Kacar,Y.,
Landry,B., Hubert,N., Laforest,M., Landry,J. and Ligonde,A.
TITLE Development of EST Resources and New Genetic Markers for California
Citrus - Washington Navel Orange Callus - UCRCS08-2
JOURNAL Unpublished (2004)
COMMENT Contact: Timothy J. Close
Department of Botany & Plant Sciences
University of California

Riverside, CA 92521-0124, USA
 Tel: 909-787-3318
 Fax: 909-787-4437
 Email: timothy.close@ucr.edu
 Seq primer: T3.
 Location/Qualifiers
 1..749
 /organism="Citrus sinensis"
 /mol_type="mRNA"
 /cultivar="Washington navel"
 /db_xref="taxon:2711"
 /clone="UCRCS08-50804-C8-1-5-g"
 /tissue_type="Callus"
 /dev_stage="Embryogenic and embryoid"
 /lab_host="E. coli TUC121"
 /clone_idb="Parent Washington Navel Orange Callus CDNA Library UCRCS08-2"
 /note="Vector: Lambda Uni-ZAP XR, excised phagemid; Site_1: EcoRI; Site_2: XhoI; Parent Washington navel orange embryogenic callus was established from undeveloped ovules of ca. 10 mm diameter young fruits under open pollination on Murashige-Skoog medium at 25°C with 16 h light in a tissue culture room. Embryogenic callus, globular and heart stage embryos were pooled in approximately equal portions in RNAlater (Ambion), then RNA was extracted using TRIzol Reagent (Invitrogen). Poly(A) RNA was purified from 500 microgram of total RNA using Qiagen Oligotex. A primary cDNA library was produced using a lambda ZAP XR cDNA Synthesis Kit (Stratagene). These steps were performed by Xinrong Ye (Roose lab, UC Riverside). One million pfu from the primary library were mass excised to produce a phagemid population by Raymond Fenton (Close lab, UC Riverside). Phagemids were plated, plasmid DNA purified, cDNA clones archived, and DNA sequences determined bi-directionally using an ABI3730 at DNA Landmarks (Landry, Hubert, Laforest, Landry, Ligonde). Chromatogram files were downloaded by PTP to UC Riverside (by Close), then processed at UC Riverside (by Manamker, Close lab), using the Harvest pipeline (http://harvest.ucr.edu) to remove vector and cloning oligo sequences and various contaminants, and to trim to a high quality region. Sequences that retained a phred 17 region of at least 100 bases were assembled, then chimeras were removed following manual inspection of assemblies (Close, Roese, Federici, Manamker, Lyon, Ye, Jang, Collin, Kacar, Ikeda, Quintilio). Sequences that survived all removal steps were submitted to Genbank."

ORIGIN

Query Match 28.6%; Score 325; DB 8; Length 749;
 Best Local Similarity 66.7%; Pred. No. 5.8e-71;
 Matches 498; Conservative 0; Mismatches 240; Indels 9; Gaps 2;
 127 TGCATAATACACCGCTTCTAGAGCTTCAACCGTCCGCAAGACGCTGGTGGCT 186
 12 TGCACAGTACAGCTTCACTCCAGTCCGCCACCAAGCACTTTGATCGGACGCT 71
 187 ACGCCGCTGAGGAGGAGATATCCGCTGTGATGCTCTCCATGGTTACTTCTAC 246
 72 TCTGATGATGCTGAGGAGAAATCCAGTACTGATTTTACTCCACGGTTACGTTCTTC 131
 247 AACTCTCTTATCTACGCTTATGCTGATGCTCTTCTCATGGCTTCACTCATCGCT 306
 132 AACTCTTCTTACCTCAGCTTATCTTACAGTTGCTCTCATGGCTTCACTTATACGCT 191
 307 CCTCAGTATATAGTATCGCGAGACGACACATGATGATTAATCAACGGCGAG 366
 192 CCTCAGTATATAGTATGCTGAGACGACATGCTGATTAATCACTTCACTGCA 251
 367 ATTATGATTTGTTATCAGTAGAATTATCACTTTCTTCAGCGCAAGTACCAAC 426
 252 ATCACAAAATGTTATCTGAAGACCTCGGCACTTCTTCCACCAACATGTTGGCCAAAT 311

427 CTATCCAAATTTGCTCTTCGCGCCATAGCCGCGTGCGAAACCGCGTTGCGGTGCC 486
 312 CTAGCAGAGCTACCCCTGCTGCTCATAGTCGAGGAGGCAAGCGCTGCTTGCATAGCT 371
 487 TTAAGAAATTTGGGTAATCTCGAATCTAAGATCTCGACATTTGCTGATAGATCCA 546
 372 CTTAAGAAAGGGG-----CCACTACTTTAAATATTCAGCATTAATGCTGTAACCTT 425
 547 GTGATGGAACAGGAAAGGAAACAAACCCCTCTCCGCTGTGCTTACCTTCCAAAC 606
 426 GTTATGAGGAATGGAACAAAGGAAACAAACCTCTCCACCGGTACTAATTACCTTCAT 485
 607 TCAATTTGACCTTAACAAACCGCTTATCTTGTGATGCTGCTGGGCTTGGTAAACCGCT 666
 486 TCAATTTGATCTTGTGT--ATGCCGGTAATGTTATGTTTCAAGCGCTTAAAGTAA 542
 667 CGAACCCTATTTTCCACCGGTGACCTCCGCGAGTGAATACCGAGGCTTTCCG 726
 543 AAAAACCTTCTGTTCTCTCTTGTGCTCGAAAGGATTAACCAAAAGCTTCTTAC 602
 727 GAATGTCAAGTCCACATGATGCTTCTGTTGCGAAGATTAATGAGCATTTGACATGCTT 786
 603 GAATGTGAGACACCGGCTGTCTATTTTGTGTTAAGATTAATGATGATGATGTTG 662
 787 GATGATGATTCAAAAGGATTAAGGAAAGATTTTATTTGTTTGTGAAGATGTTGA 846
 663 GACGATGATTAATTAAGGAGATCAGGGGGAACAAACATACGTCTGTGAAGATGTTAAG 722
 847 GAGAGGAGCAATGAGGAGATTCGTT 873
 723 TCCAGGAACTTAAGAGAGTTCATT 749
 RESULT 11
 CR397401
 LOCUS
 DEFINITION
 CR397401 504 bp DNA linear GSS 02-MAY-2004
 genomic survey sequence.
 CR397401
 CR397401.1 GI:46938129
 GSS.
 Arabidopsis thaliana (thale cress)
 Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 1
 L.Y., Rosso, M.G., Strizhov, N., Viehoveer, P. and Weissshaar, B.
 GABI-Kat SimpleSearch: a flanking sequence tag (FST) database for
 the identification of T-DNA insertion mutants in Arabidopsis
 thaliana
 Bioinformatics 19 (11), 1441-1442 (2003)
 2
 12874060
 JOURNAL
 PUBMED
 12874060
 AUTHORS
 Rosso, M.G., L.Y., Strizhov, N., Reiss, B., Dekker, K. and
 Weissshaar, B.
 TITLE
 An Arabidopsis thaliana T-DNA mutagenized population (GABI-Kat) for
 flanking sequence tag-based reverse genetics
 JOURNAL
 PUBMED
 14756321
 3
 Strizhov, N., L.Y., Rosso, M.G., Viehoveer, P., Dekker, K.A. and
 Weissshaar, B.
 High-throughput generation of sequence indexes from T-DNA
 mutagenized Arabidopsis thaliana lines
 Biotechniques 35 (6), 1164-1168 (2003)
 4
 14682050
 JOURNAL
 PUBMED
 14682050
 4 (bases 1 to 504)
 Strizhov, N., Rosso, M.G., L.Y. and Weissshaar, B.
 Direct Subcloning
 Submitted (01-MAY-2004) Weissshaar B., Max-Planck-Institut fuer
 Zuechtungsforchung, Carl-von-Linne-Weg 10, Koeln, 50829, Germany
 This sequence has been recovered from the left border of the T-DNA.

It indicates an insertion close to or within gene At5g43860. Details on the protocols used for generation of the sequence are described in References 1-3. The sequences are generated at the MPI for Plant Breeding Research in the context of the GABI-Kat project. GABI-Kat is part of the German Plant Genomics program designated 'GABI'. Information on line availability can be found at: <http://www.mpiz-koeln.mpg.de/GABI-Kat/>.

FEATURES

SOURCE

Location/Qualifiers

1..504

/organism="Arabidopsis thaliana"

/mol_type="genomic DNA"

/db_xref="taxon:3702"

/clone="GK-461A10-026076"

/clone_lib="Arabidopsis thaliana T-DNA insertion lines"

/ecotype="Col-0"

/note="PCR was performed on DNA from Arabidopsis thaliana plants (T1) which were transformed with the T-DNA from vector PAC161 (GenBank accession number: AJ537514). The lines contain one or more T-DNA insertions. The DNA fragment(s) resulting from the PCR were directly sequenced to determine the genomic sequence flanking the insertion. T-DNA derived sequences were removed."

ORIGIN

Query Match 27.5%; Score 312.4; DB 11; Length 504;
Best Local Similarity 99.7%; Pred. No. 8.3e-68;
Matches 313; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Oy 1 AAAAAAAAAAGAAAAAACTAATAAACAAAAAAAATGCTCTTTTCATCA 60
Db 115 AAAAAAAAAAGAAAAAACTAATAAACAAAAAAAATGCTCTTTTCATCA 174
Oy 61 AGAAACGCTTGAAGATGCAAAATCAAAATCTTAACTTGAATCATCATCT 120
Db 175 AGAAACGCTTGAAGATGCAAAATCAAAATCTTAACTTGAATCATCATCT 234
Oy 121 CGTTGCTCAAAATTAACCGCTCTTCTAGAGCTTCAACCGTCCGCCAAGCAGCTGTG 180
Db 235 CGTTGCTCAAAATTAACCGCTCTTCTAGAGCTTCAACCGTCTCCGCCAAGCAGCTGTG 294
Oy 181 GTGGCTACGCGGTGAGAGAGAGATTATCCGCTGTGATGCTCTTCATGTTACCTT 240
Db 295 GTGGCTACGCGGTGAGAGAGAGATTATCCGCTGTGATGCTCTTCATGTTACCTT 354
Oy 241 CTTAACAATCTCTTATTTCTAGCTTATGTTGATGCTCTTCATGAGGCTCATCTC 300
Db 355 CTTAACAATCTCTTATTTCTAGCTTATGTTGATGCTCTTCATGAGGCTCATCTC 414
Oy 301 ATCGCTCCTCAGTT 314
Db 415 ATCGCTCCTCAGTT 428
```

RESULT 12 317 bp mRNA linear EST 23-FEB-2004
AV556714/c Arabidopsis thaliana green siliques Columbia Arabidopsis
LOCUS AV556714 CDNA clone S0050d08F 3', mRNA sequence.
DEFINITION thaliana
ACCESSION AV556714.1 GI:8728129
VERSION AV556714.1
KEYWORDS EST.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana

REFERENCE
AUTHORS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
TITLE Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
1 (bases 1 to 317) rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE
AUTHORS Asamizu, E., Nakamura, Y., Sato, S. and Tabata, S.
TITLE A large scale analysis of cDNA in Arabidopsis thaliana: Generation
of 12,028 non-redundant expressed sequence tags from normalized and
size-selected cDNA libraries

JOURNAL DNA Res. 7 (3), 175-180 (2000)
PUBMED 10907847

COMMENT

Contact: Erika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizu@kazusa.or.jp, URL: <http://www.kazusa.or.jp/en/plant/>.

FEATURES

SOURCE

Location/Qualifiers

1..317

/organism="Arabidopsis thaliana"

/mol_type="mRNA"

/ecotype="Columbia"

/db_xref="taxon:3702"

/clone="S0050d08F"

/clone_lib="green siliques"

/note="Vector: BluescriptII SK-, Site_1: EcoRI; Site_2:
XhoI"

ORIGIN

Query Match 27.2%; Score 308.6; DB 1; Length 317;
Best Local Similarity 98.7%; Pred. No. 6.9e-67;
Matches 311; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
Oy 797 CAAAGGATTAAGAGGAGAGATTCTATTGTTGTAAGAATGTGAAGAGAGAC 856
Db 317 CAAAGGATTAAGAGGAGAGAGATTCTATTGTTGTAAGAATGTGAAGAGAGAC 258
Oy 857 CAATGAGAGATTGCTGTGAGACTTGTATCATTTTGAAGCTTATTTGAAGAG 916
Db 257 CAATGAGAGATTGCTGTGAGACTTGTATCATTTTGAAGCTTATTTGAAGAG 198
Oy 917 ATGATCGTAATTAGTTAATGATCAAGATGGGTGTCAGAGATGTTCCGTTGAATTC 976
Db 197 ATGATCGTAATTAGTTAATGATCAAGATGGGTGTCAGAGATGTTCCGTTGAATTC 138
Oy 977 AAGAGTTGAGTTATCATGATTAACATTAATTTTCTTAAGGCGCTGTTTCTATTGT 1036
Db 137 AAGAGTTGAGTTATCATGATTAACATTAATTTTCTTAAGGCGCTGTTTCTATTGT 78
Oy 1037 CAATATCATGCTTTGTTGCTTATGTTTACAACTTATATTTGACAACCTTTTAA 1096
Db 77 CAATATCATGCTTTGTTGCTTATGTTTACAACTTATATTTGACAACCTTTTAA 18
Oy 1097 TCACCTCTTTGCTTA 1111
Db 17 TCACCTCTTTGCTTA 3
```

RESULT 13 542 bp mRNA linear EST 14-MAY-2005
CN736070 Brassica napus 26RDBNT UP_011_F07_20JAN2004_053 Brassica napus 26RDBNT Brassica
LOCUS CN736070
DEFINITION napus cDNA 5', mRNA sequence.
ACCESSION CN736070.1 GI:65293887
VERSION CN736070
KEYWORDS EST.
SOURCE Brassica napus (rape)
ORGANISM Brassica napus

REFERENCE
AUTHORS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
TITLE Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
1 (bases 1 to 542) rosids; eurosids II; Brassicales; Brassicaceae; Brassica.

REFERENCE
AUTHORS Xiang, D., Venglat, P., Keller, G., Selvaraj, G. and Datta, R.
TITLE Gene Expression Patterns during Brassica kyoctic Embryogenesis
JOURNAL Unpublished (2004)
COMMENT Contact: Raju Datta
Molecular and Developmental Genetics
National Research Council of Canada
Plant Biotechnology Institute, 110 Gymnasium Place, Saskatoon,
Saskatchewan, S7N 0W9, Canada
Tel: 306 975 5267
Fax: 306 975 4859
Email: Raju.Datta@nrc-cnrc.gc.ca
High quality sequence stop: 542.

Tue Mar 21 06:16:23 2006

us-10-634-548-19.rst

Page 11

Db 123 GACCTCGTGAATTATAGCTATTAAGA 95

Search completed: March 20, 2006, 16:53:01
Job time : 5146.9 secs

100% 100% 100% 100% 100%

THIS PAGE LEFT BLANK

[illegible]

Db	1081	TGTACAACTCTTTAGTCACCTCTTGGCTTACAAAAA	1135
RESULT 2			
ADJ98181			
ID	ADJ98181	standard; DNA; 1135 BP.	
XX			
AC	ADJ98181;		
XX			
DT	06-MAY-2004	(first entry)	
XX			
DE	Thale cress chlorophyllase 2 DNA.		
XX			
KW	phycol kinase; tocopherol biosynthesis; plant; drought resistance;		
XX	thale cress; ds; chlorophyllase 2.		
OS	Arabidopsis thaliana.		
XX			
PN	WO2004013312-A2.		
XX			
PD	12-FEB-2004.		
XX			
PF	05-AUG-2003; 2003WO-US025276.		
XX			
PR	05-AUG-2002; 2002US-0400689P.		
XX	05-AUG-2003; 2003US-00634548.		
PA	(MONS) MONSANTO TECHNOLOGY LLC.		
XX			
PI	Norris SR, Lincoln K, Abad MS, Ellers R, Harteuiker KK;		
PI	Hirschberg J, Karunananda B, Moshiri F, Stein JC, Valentin HB;		
PI	Venkatesh TV;		
XX			
DR	WPI, 2004-157125/15.		
XX			
PT	New phycol kinase polynucleotides, useful in mediating tocopherol		
XX	biosynthesis and in producing plants with increased drought resistance.		
PS	Claim 23; SEQ ID NO 19; 189bp; English.		
CC	The invention relates to a novel substantially purified nucleic acid		
CC	molecule encoding a phycol kinase. The nucleic acid molecules and		
CC	polypeptides of the invention may be useful in mediating tocopherol		
CC	biosynthesis and in producing plants with increased drought resistance.		
CC	The current sequence is that of the thale cress chlorophyllase 2 DNA of		
CC	the invention.		
XX			
SEQ	Sequence 1135 BP; 331 A; 233 C; 247 G; 324 T; 0 U; 0 Other;		
Query Match	100.0%; Score 1135; DB 12; Length 1135;		
Best Local Similarity	100.0%; Pred. No. 9.4e-283; Indels 0; Gaps 0;		
Matches 1135;	Conservative 0; Mismatches 0;		
QY	1 AAAAAAAAAAGAAAAAAAGAAAAAACTAATAAAGAACAAAAAAATGTCCTTTTCATCA 60		
DB	1 AAAAAAAAAAGAAAAAAAGAAAAAACTAATAAAGAACAAAAAAATGTCCTTTTCATCA 60		
QY	61 AGAAACGGCTTGAAGATGGCAATATCAATATCTTTAACTTGGATCATCATCT 120		
DB	61 AGAAACGGCTTGAAGATGGCAATATCAATATCTTTAACTTGGATCATCATCT 120		
QY	121 CGTGGCTCAAAATATACACCGTCTTCTAGACCTTACCGTCTCGCCAAAGAGCTGTG 180		
DB	121 CGTGGCTCAAAATATACACCGTCTTCTAGACCTTACCGTCTCGCCAAAGAGCTGTG 180		
QY	181 GTGGCTACGCGGTGGAGAGAGATTATCCGGTGTGATGCTTCCATGTACCTT 240		
DB	181 GTGGCTACGCGGTGGAGAGAGATTATCCGGTGTGATGCTTCCATGTACCTT 240		
QY	241 CTCTACAACTCTTATCTTACCTTATGTGATGCTCTTTCAGAGCTTACCTC 300		
DB	241 CTCTACAACTCTTATCTTACCTTATGTGATGCTCTTTCAGAGCTTACCTC 300		

```

Qy 301 ATCGCTCTCATATTATAGTATGCGGAGACCAACATAGATGATTAATCAACG 360
Db 301 ATCGCTCTCATATTATAGTATGCGGAGACCAACATAGATGATTAATCAACG 360
Qy 361 GCGAGATTAATGATTTGGTTATCAGTAGAGACTTAATCACTTTCTCCAGCCAAATACA 420
Db 361 GCGAGATTAATGATTTGGTTATCAGTAGAGACTTAATCACTTTCTCCAGCCAAATACA 420
Qy 421 CCAAACTATCAAAATTTGCTCTCCGCGCATAGCCGCGTGGCAAAACCGCTTTGCG 480
Db 421 CCAAACTATCAAAATTTGCTCTCCGCGCATAGCCGCGTGGCAAAACCGCTTTGCG 480
Qy 481 GTGCGCTTAAGAATTTGGTACTCTCCGATCTTAAGATCTGACATTTGATCGATATA 540
Db 481 GTGCGCTTAAGAATTTGGTACTCTCCGATCTTAAGATCTGACATTTGATCGATATA 540
Qy 541 GATCAGTCATGAGAACAGGAAAGGAAACAAACCCCTCCGCGTGGTGGCTTACCTT 600
Db 541 GATCAGTCATGAGAACAGGAAAGGAAACAAACCCCTCCGCGTGGTGGCTTACCTT 600
Qy 601 CCAAACTCATTTGACCTTAGACAAAACGCTTAATCTGTGATCGGTTGCGGCTTGTA 660
Db 601 CCAAACTCATTTGACCTTAGACAAAACGCTTAATCTGTGATCGGTTGCGGCTTGTA 660
Qy 661 ACCGCTGGAAACCAATTAATCCACCGTGTGACCTCCGCGAGTAATCAACGAGATTC 720
Db 661 ACCGCTGGAAACCAATTAATCCACCGTGTGACCTCCGCGAGTAATCAACGAGATTC 720
Qy 721 TTTCGGGATGTCAAGGTCCAGCATGTCATTTGTTGCAAGATTAATGCGATTTGAC 780
Db 721 TTTCGGGATGTCAAGGTCCAGCATGTCATTTGTTGCAAGATTAATGCGATTTGAC 780
Qy 781 ATGCTTGATGATGATCAAAAGGATTAAGAGGAAAGTCTTAATGTTGTTGTAAGAT 840
Db 781 ATGCTTGATGATGATCAAAAGGATTAAGAGGAAAGTCTTAATGTTGTTGTAAGAT 840
Qy 841 GGTAGAGAGGAGAACCAATGAGAGATCGTGTGATGATGTTGTAATCAATTTTGAAG 900
Db 841 GGTAGAGAGGAGAACCAATGAGAGATCGTGTGATGATGTTGTAATCAATTTTGAAG 900
Qy 901 GCTTAATTTGAGAGATGATCGTGAATTAAGTAAGATCAAAAGATGGGTCAAGAGAT 960
Db 901 GCTTAATTTGAGAGATGATCGTGAATTAAGTAAGATCAAAAGATGGGTCAAGAGAT 960
Qy 961 GTTCCCGTGAATTCAGAGTTGAGGTTATCATGTAAACATTAAGTTTCTTTAGGGG 1020
Db 961 GTTCCCGTGAATTCAGAGTTGAGGTTATCATGTAAACATTAAGTTTCTTTAGGGG 1020
Qy 1021 CTGCTTTTCTAATGTCAATCATCAAGCTTTGTTGCTTAATGTTTACAACTTAAT 1080
Db 1021 CTGCTTTTCTAATGTCAATCATCAAGCTTTGTTGCTTAATGTTTACAACTTAAT 1080
Qy 1081 TGTACAACCTTTTAAGTCACTCTTTGCTTACAAAAAATTTTAAAAA 1135
Db 1081 TGTACAACCTTTTAAGTCACTCTTTGCTTACAAAAAATTTTAAAAA 1135

```

RESULT 3

```

AAD38075
ID AAD38075 standard; cDNA, 1216 BP.
XX
AC AAD38075:
XX
XX 10-SEP-2002 (first entry)
XX
XX Grape chlorophyllase cDNA from vrlc.pk008.021:fls clone.
XX
XX Grape chlorophyllase; chlorophyll degradation; plant cell senescence;
XX enzyme; gene; ss.
XX
XX vrlc sp.
XX
XX Key Location/Qualifiers
FH

```

```

FT CDS
FT 3..938
FT /*tag= a
FT /product= "Grape chlorophyllase protein from
FT vrlc.pk008.021:fls clone"
FT /EC number= "3.1.1.14"
FT /note= "CDS does not include start codon"
FT /partial
FT
FT MO200229022-A2.
FT
FT 11-APR-2002.
FT
FT 04-OCT-2001; 2001WO-US031059.
FT
FT 05-OCT-2000; 2000US-0238161P.
FT
FT (DUPO ) DU PONT DE NEMOURS & CO E. I.
FT
FT Cahoon EB, Cahoon RE, Thorpe CJ;
FT
FT WPI; 2002-444102/47.
FT
FT P-PSDB; AAE23777.
FT
FT An isolated polynucleotide encoding a plant chlorophyllase polypeptide,
FT used to produce transformed plants that have controlled induction or
FT postponement of senescence.
FT
FT Claim 5; Page 48-49; 69pp; English.
FT
FT The invention relates to plant chlorophyllase (EC 3.1.1.14) proteins and
FT polynucleotides encoding such proteins. Chlorophyllase enzyme is involved
FT in chlorophyll degradation in plant cell senescence. Sequences of the
FT invention are used to produce a plant by transforming a plant cell with
FT chlorophyllase and regenerating a plant from the transformed plant cell.
FT They may also be used to transform cells. The plants that are produced
FT can have economical importance as they can allow for the controlled
FT induction or postponement of senescence. The present sequence is grape
FT chlorophyllase cDNA from vrlc.pk008.021:fls clone
FT
FT Sequence 1216 BP; 331 A; 275 C; 256 G; 354 T; 0 U; 0 Other;
FT
FT Query Match 32.4%; Score 367.4; DB 6; Length 1216;
FT Best Local Similarity 63.9%; Pred. No. 2.2e-84;
FT Matches 608; Conservative 0; Mismatches 331; Indels 12; Gaps 3;
FT
FT 50 CTTCTTCAATCAAGAAACGCTTTGAAGATGCAAAATCAATCAATCTTTAACTTGG 109
FT 1 CTTCTTCAATGTTCCACACAGTGTGCAAAATGTTTGAAGATGGAAGACATCAAGTGC 60
FT
FT 110 ACTCATCATCTCGTGTGCAAAATTAACACCGTCTTGAAGCTTCAACCTTCCGCAA 169
FT
FT 61 TTCTAAGAGCTGAACAGGACATTCGACCAACCAAGTCATCTTCTGTTCCCTTCAC 120
FT
FT 170 AGCAGCTGTTGGGTGCTAGCGCGTGAAGAGAGATTAATCGGCTGTGATGCTCTCC 229
FT 121 TGCACCTGTGATTTGCTAGCCCTCTGAAGCAGGAGATTTCCGCTGCTCTCTCTTC 180
FT
FT 230 ATGTTAATCTTCTTCAATCTCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATG 289
FT 181 ATGTTAATCTTCTTCAATCTTCTTCAATCTTCTTCAATCTTCAATCTTCAATCTTCA 240
FT
FT 290 GCTTCAATCTTCAATCTTCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTT 349
FT 241 GTTCAATCTTCTTCAATCTTCTTCAATCTTCTTCAATCTTCTTCAATCTTCTTCAAT 300
FT
FT 350 TTAATCAACGGGAGGATTAAGATGATTAAGATTAAGATTAAGATTAAGATTAAGATTA 409
FT 301 TCAAGTCCGAGTGTCTTAAACAAATTTGTTATCAAAAGACTCCATGACTTAATCTTCT 360
FT
FT 410 CGCAATTAACCAAACTATCAAAATTTGCTCTCCGCGCATAGCCGCGTGGCAAA 469
FT 361 CCATGTTGGGCCAAATTTTAAGCAAACTAGACTTCCGCGCATATGCTGTGAGGCAAAA 420

```

```
OY 470 CCGCGTTGCGGTCGCTTAAAGAAATTTGGGTACTCCGATCTAATAAGATCTGCAAT 529
DB 421 CTGCTTTTGTCTAGACTGAGAAA-----GCATCCACTTCTTGAAATTTTCAGCTT 474
OY 530 TGATCGGTATGATGACAGTCCGATGAGACAGGAAAGGAAACAAACCCCTCCGAGT 589
DB 475 TGATGAGCATGAGACCCGGTGTGATGAAATGAGACAAAGGAAACAAACCCCTCCAGCGTAC 534
OY 590 TGGCTTACCTTCCAACTCATTTGACCTAGACAAAACCCCTATCTATCTGTGATCGGTTGG 649
DB 535 TCACCTATGTCCTCATTTCAATTTGATCTAGAC--ATGGCAGTGAATGGTAAATGGTTGG 591
OY 650 GCGTTGTGAAACCGCTCGAACCCTATTTCCACCGTGTGACCTCCGAGTGAATC 709
DB 592 GTTTGGGTGAAGTAAAGAAACCTCTGTTCCCTCTGTGCCCCCAAGGCGGTAACC 651
OY 710 ACCGAGTCTCTTCCGGAATGTCAAGGTCCAGATGACATTCCTTCCGAAAGATTATG 769
DB 652 ATGAGGACTTCTTTAAAGAAATGCGGTGAACCGCTTGTATTTCTTCCAAAGCACTATG 711
OY 770 GGCATTTGAGCATGCTTGATGATGATACAAAAGGATTAGAGAGATTTCTTATGTT 829
DB 712 GCACATCTTGACATGCTAGACGATGAGACTTAATGAAATTAAGAGGAAAGCTACACATTTGT 771
OY 830 TGTGTAAAGATGATGAAAGAGAGACCAATGAGAGATTCGTTGTGACTTGTGTAT 889
DB 772 TGTGTAAAGATGAGAGCTAGAGAACCCATGAGAGATTTGTGTGAGGATTTGTGATG 831
OY 890 CATTTTGAAGCTTATTTTGAAGAGATGATGTTGAATTAAGTCAAAAGTGGT 949
DB 832 CATTTATGAAGCTTATTTTGAAGAGGAGATACAGAGATCTAATCTCCATTTGAGATGG- 890
OY 950 GTCAAGAGAGTTCCTCCGTTGAATTCAGAGTTTGAAGTTATCATGTTAA 1000
DB 891 --CATGTAAGTCAAGAGTGAAGCTTCAAAATGTTGATTTCTCGTGTGA 939
```

RESULT 4
AAD38080
ID AAD38080 standard; cDNA, 1444 BP.

```
XX AAD38080;  
AC 07-AUG-2003 (revised)  
DT 10-SEP-2002 (first entry)  
XX Tulip chlorophyllase cDNA from etp1c.pk005.d16:fls clone.  
DE Tulip chlorophyllase; chlorophyll degradation; plant cell senescence;  
KM enzyme; gene; ss.  
XX Tulipa gesneriana.  
OS  
XX  
FH Key Location/Qualifiers  
FT CDS 97..1107  
FT /*tag= a  
FT /product= "Tulip chlorophyllase protein from  
FT etp1c.pk005.d16:fls clone"  
FT /EC_number= "3.1.1.14"  
XX  
XX WO200229022-A2.  
XX  
XX 11-APR-2002.  
XX  
XX 04-OCT-2001; 2001WO-US031059.  
XX  
XX 05-OCT-2000; 2000US-0238161P.  
XX  
XX (DUPO ) DU PONT DE NEMOURS & CO E I.  
XX  
XX Cahoon EB, Cahoon RE, Thorpe CJ;  
XX  
XX WPI, 2002-44102/47.
```

DR P-PSDB; AAE23782.
XX
PT An isolated polynucleotide encoding a plant chlorophyllase polypeptide,
FT used to produce transformed plants that have controlled induction or
PT postponement of senescence.
XX
PS Claim 5; Page 56-57; 69pp; English.
XX
CC The invention relates to plant chlorophyllase (EC 3.1.1.14) proteins and
CC polynucleotides encoding such proteins. Chlorophyllase enzyme is involved
CC in chlorophyll degradation in plant cell senescence. Sequences of the
CC invention are used to produce a plant by transforming a plant cell with
CC chlorophyllase and regenerating a plant from the transformed plant cell.
CC They may also be used to transform cells. The plants that are produced
CC can have economical importance as they can allow for the controlled
CC induction or postponement of senescence. The present sequence is tulip
CC chlorophyllase cDNA from etp1c.pk005.d16:fls clone. (Updated on 07-AUG-
CC 2003 to correct OS field.)
XX
SQ Sequence 1444 BP; 460 A; 295 C; 298 G; 391 T; 0 U; 0 Other;

Query Match 22.6%; Score 256.8; DB 6; Length 1444;
Best Local Similarity 61.1%; Pred. No. 8.9e-56;
Matches 474; Conservative 0; Mismatches 287; Indels 15; Gaps 3;

```
OY 154 TCACGATCTCCGCAAGCAGCTGTTGCTACCGCGGAGGAGAGATTTATCCG 213
DB 199 TCCGACACACCTCGGAAAGCCACCACTATGCAACTCATTTGAGAAAGGAGTACCA 258
OY 214 GTGTGATGCTCTCCATGATCTTCTTACCACTCTTATTTCTACGTTATGTTG 273
DB 259 ACGCTTACTCTTCAATGATTCATGCTTCAACAACCTTTTATCTCTGAGCTTATCCAG 318
OY 274 CATCTCTTCTATGCTTCTATCTCATGCTTCTCATGTTATATATATGCGCGGACCA 333
DB 319 CACATTCGATCCCATGCTTATGTTGTGTACCTCAAGTTATATCTTGTAGTACATGT 378
OY 334 GACCAATGATGATGATTAATCAACGCGGAGATTATGATGTTATCAGTAGGACTT 393
DB 379 GATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 438
OY 394 AATCACTTTCTTCCAGGCAAGTAACCAACCACTATCCAAATTTGCGCTCTCGGCAAT 453
DB 439 CAAGATGCTCTCCCAACAAAGTCAAGACCACTTAAGAAATCTCGAGTGAAGGACAT 498
OY 454 AGCGCGGTGCAAAACCGGCTTTGCGCTTCCCTTAAGAAATTTGGGTACTCTCGAAT 513
DB 499 AGCGGTGCGGCAAAAGTGAATTTGCTTGTGCACTAGATATGCGAAG-----ACTACA 552
OY 514 CTAAAGATCTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 573
DB 553 TTAAGCTTTCAGGCGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 612
OY 574 ACCCTCTCTCGGCTGTTGCTTACCTTCAAACTCATTTGACCTAGACAAAACGCTATA 633
DB 613 ACCAATCTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 669
OY 634 CTGTGATGCTTGTGCGGCTTGTGGAACCGCTCGGAACCATATTTCCACCGTGTGCA 693
DB 670 TTAAGTTATCGGAACAGGTTTAAAGTGAAGAGGAACCTATTTGCT-----TGTGCC 723
OY 694 CCTCCGAGTGAATCAACGAGATTTCTTCCGGAATGTCAAGGTCAGACATGACATTTTC 753
DB 724 CTTAAGGAGATTAATCAACAGATTTTAAAGATGATGATGATGATGATGATGATGATGATG 783
OY 754 GTTGCAGAGATTATGCGCATTTGAGACATGCTGATGATGATGATGATGATGATGATGATG 813
DB 784 GTTGTAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 843
OY 814 AAGAGTTCTTATGTTTGTGTAAGATGATGATGATGATGATGATGATGATGATGATGATG 873
DB 844 ATTGTGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 903
```

OY 874 GTGGAGCTGTGTTATCATTTTGAAGCTTATTGAGAGGATGATCGTAAT 929
 DB 904 GTGGAGATCGTGTGTTCTTTCTGAAGCTGTCTTACAGAGTAAACAGATTT 959

RESULT 5

AAD38074
 ID AAD38074 standard; cDNA, 987 BP.

XX AAD38074;

DT 10-SEP-2002 (first entry)

DE Grape chlorophyllase cDNA from vdb1c.pk002.p19:fls clone.

KW Grape; chlorophyllase; chlorophyll degradation; plant cell senescence;

XX enzyme; gene; ss.

OS Vicia sp.

XX Key Location/Qualifiers

FT CDS

FT 1..858

FT /product= "Grape chlorophyllase protein from

FT vdb1c.pk002.p19:fls clone"

FT /EC_number= "3.1.1.14"

FT /note= "CDS does not include start codon"

FT /partial

PN W0200229022-A2.

PD 11-APR-2002.

XX 04-OCT-2001; 2001WO-US031059.

XX 05-OCT-2000; 2000US-0238161P.

XX (DUPO) DU POINT DE NEMOURS & CO E I.

XX Cahoon EB, Cahoon RE, Thorpe CJ;

XX WPI; 2002-444102/47.

XX P-PSDB; AAE23776.

PT An isolated polynucleotide encoding a plant chlorophyllase polypeptide,

PT used to produce transformed plants that have controlled induction or

PT postoponement of senescence.

XX Claim 5; Page 47; 69pp; English.

XX The invention relates to plant chlorophyllase (EC 3.1.1.14) proteins and
 CC polynucleotides encoding such proteins. Chlorophyllase enzyme is involved
 CC in chlorophyll degradation in plant cell senescence. Sequences of the
 CC invention are used to produce a plant by transforming a plant cell with
 CC chlorophyllase and regenerating a plant from the transformed plant cell.
 CC They may also be used to transform cells. The plants that are produced
 CC can have economical importance as they can allow for the controlled
 CC induction or postoponement of senescence. The present sequence is grape
 CC chlorophyllase cDNA from vdb1c.pk002.p19:fls clone

XX Sequence 987 BP; 260 A; 224 C; 218 G; 285 T; 0 U; 0 Other;

Query Match 20.9%; Score 237.4; DB 6; Length 987;

Best Local Similarity 56.3%; Pred. No. 7.9e-51;

Matches 551; Conservative 0; Mismatches 411; Indels 17; Gaps 5;

OY 158 CGTCCGCGCAAGCAAGCTGTGAGCTAGCCGCTGAGAGAGGATTAATCCGGTGG 217

DB 23 CTTCCCTCCCAAGCATGTTGATGTTACACCAACCATTCAGGAGACATACCCAGTTC 82

OY 218 TGATGCTCTTCATGATGATACCTTCTTACAACTCCTTATCTCAGCTTAATGTTGATG 277

DB 83 TCTGTTCTTCAATGAGCTTGAGCTCGGCAACACTTCTCACTCACTGCTTCAACTCA 142

OY 278 TCTCTTCATGAGCTTCACTTCATGCTGCTCTCAATTATATGATGCGCGACAGACA 337
 DB 143 TTTCTTCCATGATTCATTTGATGATGCTCTCACTGATTAACAGCACTTATCTCTCTG 202
 OY 338 CAATGATGATTAATTAACCGCGGAGATTAATGATGATGATGATGATGATGATGATG 397
 DB 203 GAATTCAGAGATCAATCAAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 262
 OY 398 ACTTCTCCAGGCAAGTAACACCAACCTATCCAAATTTGCCCTCCGCGCATAGCC 457
 DB 263 CTGCTCTCCAGAAAATGTAACACCACTTCACTCAAGCTTGTCTTTCAAGCCACAGCA 322
 OY 458 GCGGTGCAAAACCGCTTTGCGGTGCTTTAAAGAAATTTGGTACTCTCGAATCTAA 517
 DB 323 GAGGGGGAAGACAGCATTTGCTCTGCACTAGGGATGCTGATACATCC-----CTCA 376
 OY 518 AGATCTGACATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 577
 DB 377 ACTTCTAGCCCTTACTAGGACTGACCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 436
 OY 578 CTCCTCCGCTGTGCTTACCTTCAAACTCATTTGACCTAGACAAACGCTATACCTG 637
 DB 437 TTCCCAAAATCTTAACCTATGTTCTTCACTTCTTCAATCTAG---CAATCCAGTTTGG 493
 OY 638 TGATCGGTTGCGGCTTGTGAAAACCGCTCGAAACCATTAATTTCCACGCTGACCTC 697
 DB 494 TTAATCGGACGCGGCTTGTGCGATGAGCAAGAACTGCTTAACATGTCATGTGCCCCAG 553
 OY 698 CCGGATGATCAACCGAGATTTCTTGGGAATGTCAAGTCCAGCATGCGATTTGCTG 757
 DB 554 ATGAGTGAACCAATGATGATTTTTCAGTGAATTAACCTCTGTGTTCCACTGTGTGA 613
 OY 758 CGAAGATTAATGAGCATTTGACATGCTGATGATGATGATGATGATGATGATGATGATG 817
 DB 614 CTACTGAATATGATCACTTGTGACATGATGATGATGATGATGATGATGATGATGATG 673
 OY 818 GTTCTTATGTTGTTGTAAGATGTGAAG--AGAGAGACCAATGAGGATGCTGTTG 874
 DB 674 TTTGCGGTTATATCTGACAAAGTGGGAAGGCTCTGAGGACCCCATGAGGATGTGTGG 733
 OY 875 GTGACCTGTGATATCAATTTTGAAGGCTTATTTGGAAGGATGATGATGATGATGATG 934
 DB 734 GTGACCTTTTGTGCACTTCTTAAGGCTTATTTGGAAGGCTGACACTGAGATTT---TCA 790
 OY 935 AGATCAAAAGATGAGGCTGACAGAGATGTTCCGTTGAAATTCAGAGTTGAGATTATCA 994
 DB 791 AAGCATGTTGATGAACCTGATCTGCTCTGTGAAGCTTGAATCTGTTGAGTTCAATG 850
 OY 995 TGTAAACATAAGTTTTTTCTTTAGGCGCTGCTTTTCTATTTGCAATATATCAAGCTTTTG 1054
 DB 851 AAGCATTAATTCATGATTTCTTGTATGCTATTAAGAGAGTGA--CAGACGATTTTCTT 908
 OY 1055 TTGCTTATGTTTATCAAACTTATATGATGATGATGATGATGATGATGATGATGATG 1114
 DB 909 TTGCTTATGTTTATGAGAAATTAATTAAGAAATTAAGAAATTAAGAAATTAAGAAAT 968
 OY 1115 AAAAAAAAAAAAAAAAAA 1133
 DB 969 AAAAAAAAAAAAAAAAAA 987

RESULT 6

AAD38076
 ID AAD38076 standard; cDNA, 1302 BP.

XX AAD38076;

XX 10-SEP-2002 (first entry)

XX Corn chlorophyllase cDNA from ceh3c.pk001.a9:fls clone.

KW Corn; chlorophyllase; chlorophyll degradation; plant cell senescence;

KM enzyme; gene; ss.
 XX
 OS Zea mays.
 XX
 XX Key Location/Qualifiers
 FH CDS 72..1112
 FT /*tag= a
 FT /product= "Corn chlorophyllase protein from
 FT csh3c.pk001.a9:file clone"
 FT /EC_number= "3.1.1.14"
 FT
 PN MO200229022-A2.
 XX
 PD 11-APR-2002.
 XX
 PF 04-OCT-2001; 2001WO-US031059.
 XX
 PR 05-OCT-2000; 2000US-0238161P.
 XX
 PA (DUPO) DU PONT DE NEMOURS & CO E. I.
 PI Cahoon EB, Cahoon RE, Thorpe CJ;
 XX
 DR WPI; 2002-444102/47.
 XX
 DR P-PSDB; AAE23778.
 XX
 PT An isolated polynucleotide encoding a plant chlorophyllase polypeptide,
 PT used to produce transformed plants that have controlled induction or
 PT postponement of senescence.
 XX
 PS Claim 5; Page 50; 69pp; English.
 XX
 CC The invention relates to plant chlorophyllase (EC 3.1.1.14) proteins and
 CC polynucleotides encoding such proteins. Chlorophyllase enzyme is involved
 CC in chlorophyll degradation in plant cell senescence. Sequences of the
 CC invention are used to produce a plant by transforming a plant cell with
 CC chlorophyllase and regenerating a plant from the transformed plant cell.
 CC They may also be used to transform cells. The plants that are produced
 CC can have economical importance as they can allow for the controlled
 CC induction or postponement of senescence. The present sequence is corn
 CC chlorophyllase cDNA from csh3c.pk001.a9:file clone
 CC
 SQ Sequence 1302 BP; 248 A; 424 C; 407 G; 223 T; 0 U; 0 Other;
 Query Match 17.8%; Score 201.6; DB 6; Length 1302;
 Best Local Similarity 54.6%; Pred. No. 1,66-41;
 Matches 449; Conservative 0; Mismatches 364; Indels 9; Gaps 2;
 QY 143 CTTCTAGAGCTTCAACCGTCTCCGCAAGCAAGCTGTGTGCTACGCGGTGAGGAAAG 202
 DB 241 CGTCAGGCGCGCCAGCGCTGCGCGAAGCCGCTCTCTGTGCGCGCGCGGCGAGACTG 300
 QY 203 GAGATTATCCGCTGTGATGCTCTCCATGCTTACCTTCTCTCAACACTCTTCTTATTTCTC 262
 DB 301 GGGAGTACCGCGGTGATCTCTTCTTCAACGCGTACTGCGCGTCAACCTCTTCTACTCC 360
 QY 263 AGCTTATGTCATGCTCTTCTCTCATGCTTCACTCATGCGCTCTAGTTATATAGTA 322
 DB 361 AGCTTGTGAGACATCGCTCTCCATGCTTATGTGTGAGACTCAAGCTGTACACCA 420
 QY 323 TCGCCGACACAGACCAATGATGATTAATCAACGCGGAGATTATGATTTGTTAT 382
 DB 421 TATCTGGGGCGGACACCGAGATCACTACGCGGCGCGCTCATCGACTGTGCTAG 480
 QY 383 CAGTAGACTTATATCACTTTCTTCCAGCGCAAGTAAACCAACTTATCCAAATTTGCC 442
 DB 481 CCACCGGCTGCGCTCAACTCTGCACTCGCGCTCGCGCGAAGCTTAACTCAAGGTGTCCA 540
 QY 443 TCTCGGCGCATAGCGCGGTTGGCAAAACCGGTTTGGCGCTTAAAGAAATTTGGGT 502
 DB 541 TCTTCGGCCACAGTGTGGCGGGAAGTGGCGTGTGCGCTGGCGTTGGGCCACCGCAAG 600
 QY 503 ACTCTCGAATCTAAAGATCTCGACATGATCGTATATGATCACTGATGGAACAGGGA 562

DB 601 CCAAGCTGCTGTCTCTCTCTGCGCGCCGCTGCGCGGTGAGACCGGCTGAGAGGATGGCG 660
 QY 563 AAGGAAACAAACCCCTCTCTCCGCTGTGCTTACTTCCAAACTATTGACCTAGACA 622
 DB 661 TGGGCAAGCAGACACCCCGCCGATCTTCAAGGAGCAGCGCTCCCTCAGCGTG--- 717
 QY 623 AAACGCTTACTTGTGATGAGTTCGCGGCTGTGTAACCGCTCGGACCCATTATTC 682
 DB 718 GTGCCCCCGCATGTGATGAGCAGGAGGCTCGCGAGCTGCCCCGCGCTCACTGCTCC 777
 QY 683 CACCGTGTGACCTCCCGAGTGAATCACCGAAGTTCT-----TTCCGGAATGTCAAG 736
 DB 778 CGCGGTGCGGCCCCCGGCGGCTCAGTCAAGCGGCTTCTTCAAGACGACTGAGCGCGCG 837
 QY 737 GTCCAGCATGSCATTTCGTTCCGAGGATTATGAGCACTTGGACATGCTGATGATATA 796
 DB 838 CGCCAGCGTCCCACTGTGTGTGCAAGGACTACGCGCACAGCATGATGAGACGACACA 897
 QY 797 CAAGAAGGATTAGAGGGAAGATTCTTATGTTTGTGTAGAAATGTGAAAGAGAGAGAC 856
 DB 898 CGCGGCGCGCAGGAGATCTCAAGCGCACCATGTGACAGAGCGGCGGCGCAGGCGC 957
 QY 857 CAATGAGAGATTGTTGTGTGACCTTGTGTATCATTTTGAAGCTTATTTGAGAGAG 916
 DB 958 CCAATGCGCGCTTGTGTGCGCGGCGCACCGTCCGCTTCTCAAGAAATGGGTGGCTGGGG 1017
 QY 917 ATGATCTGTAATTGATTAAATCAAGATGGGTGTACGAGG 958
 DB 1018 ACGCGCGCGGATGAGACAGCATCAAGCGCGCGCGGACGAG 1059
 RESULT 7
 ID ADX51602
 XX ADX51602 standard; cDNA; 1125 BP.
 XX
 AC ADX51602;
 XX
 XX 21-APR-2005 (first entry)
 DT
 XX
 DE Plant full length insert polynucleotide seqid 26342.
 XX
 XX plant protectant; plant growth regulant; gene therapy; plant;
 KW recombinant DNA construct; physical array; plant breeding marker;
 KW cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
 KW extreme osmotic condition; pathogen tolerance; pest tolerance;
 KW growth rate; cell cycle pathway; disease resistance;
 KW galactomanan production; lignin production; plant growth regulator;
 KW yield; plant growth; plant development; seed oil; protein yield;
 KW protein content; gene; ss.
 XX
 OS Unidentified.
 XX
 XX US204034888-A1.
 PN
 PD 19-FEB-2004.
 PF
 XX 28-APR-2003; 2003US-00425114.
 XX
 PR 06-MAY-1999; 99US-00304517.
 PR 05-NOV-2001; 2001US-00985678.
 XX
 XX (LIU/J) LIU J.
 PA (ZHOU/Y) ZHOU Y.
 PA (KOVA/J) KOVALIC D K.
 PA (SCRE/J) SCREEN S E.
 PA (TABAS/J) TABASKA J E.
 PA (CAO/Y) CAO Y.
 XX
 PI Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;
 XX WPI; 2004-180133/17.
 DR

PT New recombinant DNA construct, useful for improving plant tolerance to
PT cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
PT pests, for conferring increased resistance to plant disease, or for
PT improving yield.

XX Claim 1; SEQ ID NO 26342; 15pp; English.

CC The invention describes a recombinant DNA construct comprising a
CC polynucleotide consisting of a sequence encoding an amino acid sequence
CC available in electronic form from the US patent office at
CC ftp.seqdata.uspo.gov/sequence.html?DocID:2004034888. The polynucleotide
CC of the invention are also useful in physical arrays of molecules and as
CC plant breeding markers. The recombinant DNA construct is useful for
CC improving plant tolerance to cold, heat, drought, herbicides, extreme
CC osmotic conditions, pathogens or pests, for manipulating growth rate in
CC plant cells by modification of the cell cycle pathway, for conferring
CC increased resistance to plant disease, for producing galactosemannan,
CC lignin or plant growth regulators, for increasing the rate of homologous
CC recombination in plants, for improving yield by modification of
CC photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
CC or by providing improved plant growth and development under at least one
CC stress condition or for modifying seed oil or protein yield and/or
CC content. This sequence represents a plant full length insert
CC polynucleotide that can be used in the recombinant DNA construct of the
CC invention.

XX Sequence 1125 BP; 293 A; 260 C; 274 G; 298 T; 0 U; 0 Other;

Query Match 16.2%; Score 184; DB 13; Length 1125;
Best Local Similarity 56.1%; Pred. No. 5.2e-37;
Matches 431; Conservative 0; Mismatches 325; Indels 12; Gaps 4;

QY 139 CCGTCTTCTAGAGCTTACCGCTCTCCGCAAGACGCTGTGTGCTACCGCGGTGAG 198
DB 135 CCGTCCGCAACCGATTGACCGCACCTCCAAAGCCTGTAATAATCACTCCCAACGCTC 194
QY 199 GAAGAGATTATCGGGGTGTGATCTCTCCATGCTTACCTTCTTCAACTCTCTTAT 258
DB 195 GCCGAACTTACCCCGTCTTATCTTCCATGATTTCTTCTGTAATCTTCTTAC 254
QY 259 TCTGAGCTTATGTGATGCTTCTTCTCATGCTTCACTCATGCTCTCACTTAT 318
DB 255 TCTGATGTTATTAACAGTAGCTTCTCATGCTTCACTGTTGTAAGCCACACACTGTC 314
QY 319 AGTATCCCGGACCAAGACAAATGATGATTAATCAACGCGGAGATTATGATGG 378
DB 315 AAGATTTTGGCCGCGGAGGCAAGTGAAGTGAACATGCTGGAAGATTAATCTGG 374
QY 379 TTATCACTAGACTTATTAATCACTTCTTCCAGCGCAAGTAACCAACATTAAT 438
DB 375 ACTTCGAAAAAATCTCAAGCTCACTCCCAAGCTCAGTAACGCTAATGGCACTACACC 434
QY 439 GCCCTCTCCGCGCATAGCCGCGGTGGCAAAACGCGTTGGCGCTTAAAGAAATTT 498
DB 435 GCACTCGTGGGCAATAGCCGCGGTGTAAACCCGCTTGGGCTTAAAGCAAGCC 494
QY 499 G---GTAATCTCTGAAATCTAAAGATCTGCAATGATCGTATGATTCAGTGA 555
DB 495 GCAACACTAGACCATCATCAAGTTTCAAGCTCTTGAAGAAATAGATCCAGTTGACGA 554
QY 556 ACAGGGAAGGGAACAAACCCCTCCCGGTGTGGCTTACCTTCAAACTCAATTTGAC 615
DB 555 ATCAAGCAATGCAATAGAACGATCCGAAATCTTAAAGTAACAAACGGAATCAATTTGAC 614
QY 616 CTAGACAAAACGCTTATATCTGTATCGGTTCGGGCTTGTGTAACCGCTCGAACC 675
DB 615 CTGGAATGCGGTTGCA---GTATGCGTACGGGTCCTCG--ACGGAAGATTAACATG 668
QY 676 TTATTTCCACCGGTGCACTCTCCGAGTGAATCAACGAGATTTCTTTGGGAATGCA 735
DB 669 CTGATGCAACCATGCGACCAAGCGGAAGTAACCATAGAGAGTTTATTTAGATGTAAG 728
QY 736 GGTCAAGATGGAATTTGTTGGAAGGATTAATGGGCAATTTGACATGCTTGAATGAT 795

DB 729 GCTACCAAGGGGCAATTCGGCTGCGGATTACGACATATGATGTTGACGATTAAT 788
QY 796 ACAAAAGGATTAGAGGAAGATTTCTTATTTGTTGTGAATGCTGA---GAGAGC 852
DB 789 TTGCGCGGTTTGTGCGGTTTATGCGCGGTGTATGTTATGTAAGACATTAACCAAAAG 848
QY 853 AGACCAATGAGAGATTCGTTGTGACCTTTGTTGATCATTTTGAAG 900
DB 849 AGTGAATGAGAAAGCTTTGTTGTGGAATTTGTTGCGTTCTTAAAG 896

RESULT 8

AB212353 standard; DNA; 975 BP.

AC AB212353;

DT 21-JAN-2003 (first entry)

DE Arabidopsis thaliana stress regulated gene SEQ ID NO 158.

KM Arabidopsis thaliana; plant; gene; stress; transgenic; ds.

OS Arabidopsis thaliana.

PN WO200216655-A2.

PD 28-FEB-2002.

PP 24-AUG-2001; 2001WO-US026685.

PR 24-AUG-2000; 2000US-0227866P.

PR 26-JAN-2001; 2001US-0264647P.

PR 22-JUN-2001; 2001US-0300111P.

PA (SCRI) SCRIPPS RES INST.

PI (SYGN) SYNGENTA PARTICIPATIONS AG.

PI Harper JF, Kreps J, Wang X, Zhu T;

DR WPI; 2002-304127/34.

PT Identifying a stress condition to which a plant cell has been exposed and
PT producing plants with increased tolerance to these abiotic stresses.

PS Claim 14; SEQ ID NO 158; 577bp + Sequence Listing; English.

CC The invention relates to identifying a stress condition to which a plant
CC cell has been exposed, comprising: (a) contacting nucleic acid
CC representative of expressed polynucleotides in the plant cell with an
CC array or probes representative of the plant cell genome; and (b)
CC characterizing a profile of expressed polynucleotides in the plant cell
CC characteristic of a stress response. The method is useful in the
CC production of transgenic plants, cells and seeds and in producing plants
CC with increased tolerance to abiotic stress. The present sequence is that
CC of an Arabidopsis thaliana stress regulated gene (AB212196-AB217574) used
CC in methods of the invention. Note: The sequence data for this patent is
CC not represented in the printed specification but is based on sequence
CC information supplied to Derwent by the European Patent Office

XX Sequence 975 BP; 250 A; 222 C; 256 G; 247 T; 0 U; 0 Other;

Query Match 14.4%; Score 163.4; DB 6; Length 975;
Best Local Similarity 54.3%; Pred. No. 1e-31;
Matches 420; Conservative 0; Mismatches 341; Indels 12; Gaps 4;

QY 151 GCTTACCGTCTCCGCAAGACGCTGTTGTGCTTACCGCGGTGAGAGAGATTTAT 210
DB 109 GATTCAACAGCACCGCAAAACCGGTGAGAAATCACTGTGCAACAGTCCCGGAATTTAT 168
QY 211 CCGGTGATGCTCTCCATGATGTTACCTTCTTACAACTCTTCTTATCTGACTTATG 270

```

Db      169 CCCGTCGTTTATCTTCATGCGCTTTTATCTTCGCAACTACTCTAGCTGAGCTCTT 228
Qy      271 TTGCATGCTCTCTTCATGAGCTTCACTCTCATCGCTCTCACTTATATAGTATGCCGGA 330
Db      229 AACCATCGCTTCCGATGCTTACATCTCTTGAGCCCAAGTTGTGAAAATTATTCGCG 288
Qy      331 CCAGACAAATGATGATTAATCAACGCGGAGATTTATGATTTGGTTTCACTAGAGA 390
Db      289 CCGGAGGGGCAAGTGGAACTGACATGCTGGAAGTGTGATTAACCTGGGCACTCGAAAC 348
Qy      391 CTTAATCACTTCTTCAGCGCAAGTAAACCAAACTATCCAAATTTCCCTCTCCGCG 450
Db      349 CTCAAAGCTCACTCACTCACTTCGCTTAATGCTAATGAAAAATACACTCTGCTGGGC 408
Qy      451 CATACCCCGGCTGGCAAAACCCGCTTTGCGCTGCTTAAAGAAATTG--GGTACTCC 507
Db      409 CACACCCCGGCTGGGAAAACCGCGTTTCGCTGCTAGGCGCAAGCCGCAACATTAGAC 468
Qy      508 TCGAATCTAAAGATCTCGACATTGATCGGTATAGATCCAGTCGATGGAACAGGGAAAGG 567
Db      469 CCATCCATCAGCTTTTCAGCTCTAATAGGAATGATCCAGTCGCAAGAACTAACAAATAC 528
Qy      568 AAACAACCCCTCTCTCCGCTGCTTACCTTCGCAACTCATTTGACCTAGACAAAACG 627
Db      529 ATTGAACCGATCCGCATATCTTAACGTAATAACCGAATCTTTGAGCTGACATACCG 588
Qy      628 CCTATACCTTGATGATGCTTCGCGGCTTGTGAAACCGCTCGGAACCCATATATCCACCG 687
Db      589 GTTGACGAGTG--GGAACCGGACTCG--ACGAAGTGAACACGTATGACCA 642
Qy      688 TGTGACCTCCCGAGTGAATCAACGAGAGTTCTTTCGGGAATGTCAAGTCCAGCATAG 747
Db      643 TCGGACCAACGAGATTAACCATAGAGATTTTCAAAAGTGTAAAGCCACGAAACCC 702
Qy      748 CATTTGCTGCGAAGATTAAGGCGATTTGACATGCTTATGATGATTAACAAAGGAGT 807
Db      703 CATTTGCTGCGTGCAGATTAACGACATATGATATGTTGACGATGATTTGCCGCTTT 762
Qy      808 AGAGGGAAGATTTTATTTGTTGTGTAAGATGTGTA--AGAGAAGACCAATAGAG 864
Db      763 GTTGAGTTTATGCGCGGTTGATGTAAGATGAGGCAAAAGAAAGCTGAGATGAGG 822
Qy      865 AGATTCGTTGAGGACTGTTGATCATTTTGAAGGCTTATTTGGAAGAGA 917
Db      823 AGCTTTGATAGTGAATGTGTGCTTTCGAAGTATAGTTTGTGGGTGA 875

RESULT 9
ADG87791
ID      ADG87791 standard; cDNA; 975 BP.
XX
AC      ADG87791;
XX
DT      22-APR-2004 (first entry)
XX
DE      A. thaliana RPP7/RPP8-upregulated pathogen infection-related gene #233.
XX
KM      Pathogen infection-related gene; plant; Peronospora parasitica;
XX      defence mechanism; RPP7; RPP8; pathogen resistance; transgenic plant;
XX      oomycete; fungus; bacterium; virus; nematode; insect; aphid; gene; ss.
XX
OS      Arabidopsis thaliana.
XX
PN      WO200222675-A2.
XX
PD      21-MAR-2002.
XX
PF      14-SEP-2001; 2001WO-US028506.
XX
PR      15-SEP-2000; 2000US-0232778P.
XX
PR      22-JUN-2001; 2001US-0300183P.
XX
PA      (SYGN ) SYNGENTA PARTICIPATIONS AG.

```

```

PA      (UYNC-) UNIV NORTH CAROLINA.
PA      (GLAZ/) GLAZEBROOK J.
PA      (WANG/) WANG X.
PA      (DANG/) DANG J L.
PA      (EULG/) EULGEM T.
PA      (ZHU/) ZHU T.
PI      Glazebrook J, Wang X, Dang J, Eulgem T, Zhu T;
XX      WPI; 2002-292409/33.
XX
PT      Novel isolated polynucleotide, useful for conveying pathogen resistance
PR      to plants, and for identifying plants infected with a pathogen.
XX
PS      Claim 3; SEQ ID NO 233; 605bp; English.
XX
CC      The invention relates to 691 Arabidopsis thaliana genes (ADG87559--
CC      ADG87571) whose expression is altered in response to pathogen infection,
CC      and to homologues of these genes from other plants or fungi, especially
CC      from maize, soybean, barley, alfalfa, sunflower, canola (oilseed rape),
CC      cotton, peanut, sorghum, tobacco, sugarbeet, rice or wheat. The
CC      expression of genes of the invention was upregulated or downregulated in
CC      Arabidopsis plants infected with the oomycete Peronospora parasitica,
CC      indicating that they play a role in defence mechanisms. The genes of the
CC      invention are regulated by RPP7 or RPP8 which act via unconventional
CC      signalling cascades, or by the RPP4-dependent pathway. The invention also
CC      relates to polypeptides encoded by the pathogen infection-related genes;
CC      promoter motifs from pathogen infection-related genes (ADG88243-ADG88327)
CC      ; expression cassettes, host cells and pathogen-resistant transgenic
CC      plants and their progeny comprising a polynucleotide of the invention;
CC      a method of identifying a plant cell infected with a pathogen. The
CC      polynucleotide sequences and methods of the invention are useful for
CC      identifying plants infected with a pathogen, and for conferring
CC      resistance to pathogens such as oomycetes, fungi, bacteria, viruses,
CC      nematodes and insects (e.g., aphids). The present sequence represents an
CC      Arabidopsis thaliana gene whose expression is altered in response to
CC      Peronospora parasitica infection. Note: The sequence data for this patent
CC      did not form part of the printed specification, but was obtained in
CC      electronic format directly from WIPO at
CC      ftp.wipo.int/pub/published_pct_sequences.
XX
SQ      Sequence 975 BP; 250 A; 222 C; 256 G; 247 T; 0 U; 0 Other;
XX
Query Match      14.4%; Score 163.4; DB 6; Length 975;
Best Local Similarity 54.3%; Pred. No. 1e-31;
Matches 420; Conservative 0; Mismatches 341; Indels 12; Gaps 4;
Qy      151 GCTTACCGTCTCCGCAAGCAAGCTGTGTGCTACGCGGTGAGAGAGATTAAT 210
Db      109 GATTCACAGCAGCCGCAAAACGGGTGAAATCACTGTCAACAGTCCGGAACCTTAT 168
Qy      211 CCGGTGATGCTCTTCATGATGTTACCTTCTCAACAACCTCTTATTTCACTTATG 270
Db      169 CCGGTGTTTATCTTCCATAGCTTATCTTTCGCAACTCTTCTGAGAGTTCTT 228
Qy      271 TTGCATGCTCTTTCATGAGCTTATCATGCTCATGCTCTGATATATAGATGCCGGA 330
Db      229 AACCATCGCTTTCGATGCTTATCTTGTACCCCAAGTTGTGCAATTTATTCGCG 288
Qy      331 CCAGACAAATGATGATTAATCAACGCGGAGATTTATGATTTGGTTTCACTAGAGA 390
Db      289 CCGGAGGGGCAAGTGGAACTGACATGCTGGAAGTGTGATTAACCTGGGCACTCGAAAC 348
Qy      391 CTTAATCACTTCTTCAGCGCAAGTAAACCAAACTATCCAAATTTGCCCTCTCCGCG 450
Db      349 CTCAAAGCTCACTCACTCACTTCGCTTAATGCTAATGAAAAATACACTCTGCTGGGC 408
Qy      451 CATACCCCGGCTGGCAAAACCGGCTTTGCGCTGCTTAAAGAAATTG--GGTACTCC 507
Db      409 CACACCCCGGCTGGGAAAACCGCGTTTCGCTGCTAGGCGCAAGCCGCAACATTAGAC 468
Qy      508 TCGAATCTAAAGATCTCGACATTGATCGGTATAGATCCAGTCGATGGAACAGGGAAAGG 567

```

DB CCATCATACAGCTTTTACGCTTATAGGAATGATCCAGTCGAGAACTAACAAATAC 528
QY 568 AAACAACCCCTCTCCGCTGCTTACCTTCCAACTCATTTGACCTAGACAAACG 627
DB 529 ATTGAACCCGATCCGATATCTTAACGATTAACCGGAATCTTTGAGCTGACATACCG 588
QY 628 CCTATACCTTGATCGCTTCCGGGGCTTGGTGAACCCGTCGGAACCCCTTATCCCAACG 687
DB 589 GTTGACAGTGTG---GGAACCGGAACTCGG---ACCGAAGTGAACACGATATGCAACA 642
QY 688 TGTGACCTCCCGAGGATGATCAACCGAGTCTTCCGGAATGTCAGATCCAGCATGG 747
DB 643 TGGGACCAACGAGACTTAACCATGAGAGTTTACAAAGCTGTAAAGCCAGCAACCC 702
QY 748 CATTTCGTTGGGAAGATTATGGGCAATTGACATGCTTGAATGATGATACAAAGGATT 807
DB 703 CATTTCGTTGGGAGATTACGACATATGATATGTTGACGATGATTTCCCGGTTT 762
QY 808 AGAGGGAAGATTCTTATGTTTGTGTAAGAAATGCTGA---AGAGAGAGACCATGAG 864
DB 763 GTTGAGTTTATAGCCGCTTGTATGTATGTAAGAAATGGCAAAAGAAAAGCTGAGATGAG 822
QY 865 AGATTGCTGGTGGATCTGTGATCATTTTGAAGGCTTATTTGGAAGAGA 917
DB 823 AGCTTTGATGATGATGTGTGCTTCTCAAGTATAGTTTGTGGGTGA 875

RESULT 10
ADG87590
ID ADG87590 standard; cDNA, 975 BP.
XX
AC ADG87590;
XX
DT 22-APR-2004 (first entry)
XX
DE A. thaliana RPP7/RPP8-upregulated pathogen infection-related gene #32.
XX
KW Pathogen infection-related gene; plant; Peronospora parasitica;
KW defence mechanism; RPP7; RPP8; pathogen resistance; transgenic plant;
KW oomycete; fungus; bacterium; virus; nematode; insect; aphid; gene; ss.
XX
OS Arabidopsis thaliana.
XX
PN MO200222675-A2.
XX
PD 21-MAR-2002.
XX
PF 14-SEP-2001; 2001MO-US028506.
XX
PR 15-SEP-2000; 2000US-0232778P.
PR 22-JUN-2001; 2001US-0300183P.
XX
PA (SYGN) SYNGENTA PARTICIPATIONS AG.
PA (UNNC-) UNIV NORTH CAROLINA.
PA (GLAZ/) GLAZEBROOK J.
PA (WANG/) WANG X.
PA (DANG/) DANG J L.
PA (EULG/) EULGEM T.
PA (ZHUT/) ZHU T.
XX
PI Glazebrook J, Wang X, Dang J, Bulgem T, Zhu T;
XX
DR WPI; 2002-292409/33.
XX
PT Novel isolated polynucleotide, useful for conveying pathogen resistance
XX to plants, and for identifying plants infected with a pathogen.
XX
PS Claim 3; SEQ ID NO 32; 605bp; English.
XX
CC The invention relates to 691 Arabidopsis thaliana genes (ADG87559--
CC ADG87557) whose expression is altered in response to pathogen infection,
CC and to homologues of these genes from other plants or fungi, especially
CC from maize, soybean, barley, alfalfa, sunflower, canola (oilseed rape),

CC cotton, peanut, sorghum, tobacco, sugarbeet, rice or wheat. The
CC expression of genes of the invention was upregulated or downregulated in
CC Arabidopsis plants infected with the oomycete Peronospora parasitica,
CC indicating that they play a role in defence mechanisms. The genes of the
CC invention are regulated by RPP7 or RPP8 which act via unconventional
CC signalling cascades, or by the RPP4-dependent pathway. The invention also
CC relates to polypeptides encoded by the pathogen infection-related genes;
CC promoter motifs from pathogen infection-related genes (ADG88243-ADG88327)
CC ; expression cassettes, host cells and pathogen-resistant transgenic
CC plants and their progeny comprising a polynucleotide of the invention;
CC a method of identifying a plant cell infected with a pathogen. The
CC polynucleotide sequences and methods of the invention are useful for
CC identifying plants infected with a pathogen, and for conferring
CC resistance to pathogens such as oomycetes, fungi, bacteria, viruses,
CC nematodes and insects (e.g., aphids). The present sequence represents an
CC Arabidopsis thaliana gene whose expression is altered in response to
CC Peronospora parasitica infection. Note: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 975 BP; 250 A; 222 C; 256 G; 247 T; 0 U; 0 Other;
Query Match 14.4%; Score 163.4; DB 6; Length 975;
Best Local Similarity 54.3%; Pred. No. 1e-31;
Matches 420; Conservative 0; Mismatches 341; Indels 12; Gaps 4;
QY 151 GCTTACCGTCTCCGCAAGACAGCTGTGGCTACGCGGTGAGGAAGATAT 210
DB 109 GATTCAACAGACCGCCAAACCGGTGAGATCACTGTCAACAGTCCGCGAATTTAT 168
QY 211 CCGGTGTGATGCTCTCCATGTTTACCTTCTCAACCTCTTATTTCACTTATG 270
DB 169 CCGGTGTGATGCTCTCCATGTTTACCTTCTCAACCTCTTATTTCACTTATG 228
QY 271 TTGATGCTCTCTCATGAGCTTCATGCTCATGCTCTCATGATATATATATGCGCGGA 330
DB 229 AACCAATGCTCTGCAAGTTACATTTGTAGGCCACACAGTTGTCAATTTATGCGG 288
QY 331 CCAAGACAAATGATGATGATTAATCAACCGCGAGATTATGATTTGATATGATGAGA 390
DB 289 CCGGAGAGGCAATGAGATGAGACAGTCTGGAAGTGTATTAATCGGATCGGAAC 348
QY 391 CTTATATCATTTCTTCCAGGCAAGTAAACCAACCTATTCAAATTTGCCCTCCGCGC 450
DB 349 CTCAAAGCTACCTACCAATTCGCTTAATGATAAATAACCTCACTCTGTTGGC 408
QY 451 CATAGCCGCGGTGGCAAAACCGGTTGGGTGCGCTTAAGAAATTTG---GGTACTCC 507
DB 409 CACAGCCGCGGTGGAAACCGGCTTGGCGGTGCGCTAGCCATGCGCAACATTAGAC 468
QY 508 TCGATCTTAAGATCTCGATGATGATGATGATGATGATGATGATGATGATGATGATG 567
DB 469 CCATCATACAGTTTTCAGCTTATATGAAATGATGATGATGATGATGATGATGATG 528
QY 568 AAACAACCCCTCTCCGCTGCTTACCTTCCAACTCATTTGACCTAGACAAACG 627
DB 529 ATTGAACCCGATCCGATATCTTAACGATTAACCGGAATCTTTGAGCTGACATACCG 588
QY 628 CCTATACCTTGATCGCTTCCGGGGCTTGGTGAACCCGTCGGAACCCCTTATCCCAACG 687
DB 589 GTTGACAGTGTG---GGAACCGGAACTCGG---ACCGAAGTGAACACGATATGCAACA 642
QY 688 TGTGACCTCCCGAGGATGATCAACCGAGTCTTCCGGAATGTCAGATCCAGCATGG 747
DB 643 TGGGACCAACGAGACTTAACCATGAGAGTTTACAAAGCTGTAAAGCCAGCAACCC 702
QY 748 CATTTCGTTGGGAAGATTATGGGCAATTGACATGCTTGAATGATGATACAAAGGATT 807
DB 703 CATTTCGTTGGGAGATTACGACATATGATATGTTGACGATGATTTCCCGGTTT 762
QY 808 AGAGGGAAGATTCTTATGTTTGTGTAAGAAATGCTGA---AGAGAGAGACCATGAG 864

Db 763 GTTGGGTTTATGCGCGGTTGTATGTGTAAGAATGGCGAAAGAAAAGCTGAGATGAGG 822
QY 865 AGATTGCTTGGTGGAGCTTGTGTATCATTTTGAAGGCTTATTTGGAAGAGA 917
Db 823 AGCTTTGATGGTGAATGTGTGTTGCTTTCACAGTATAGTTGTGGGGTGA 875

RESULT 11

AAH41136
ID AAH41136 standard; DNA, 1188 BP.

AAH41136;
AC

DT 21-AUG-2001 (first entry)

DE Arabidopsis thaliana chlorophyllase coding sequence #1.

XX Chlorophyllase; transgenic plant; ds.

XX Arabidopsis thaliana.

PN JP2001086990-A.

PD 03-APR-2001.

PF 20-SEP-1999; 99JP-00266181.

PR 20-SEP-1999; 99JP-00266181.

PA (KAGO) KAGOME KK.

DR WPI; 2001-338421/36.

DR P-PSDB; AAB9102.

PT DNA encoding chlorophyllase, useful for producing transgenic plants.

PS Claim 3; Page 10-11; 21pp; Japanese.

CC The present sequence is a chlorophyllase protein coding sequence. The

CC chlorophyllase protein coding sequence can be used for the transformation

CC of a plant

SQ Sequence 1188 BP; 327 A; 255 C; 281 G; 325 T; 0 U; 0 Other;

Query Match 14.4%; Score 163.4; DB 5; Length 1188;

Best Local Similarity 54.3%; Pred. No. 11e-31;

Matches 420; Conservative 0; Mismatches 341; Indels 12; Gaps 4;

QY 151 GCTTCACGCTCCGCAAGAGCTGTTGGTGGCTACCGCGTGAAGAGAGATTAT 210
Db 157 GATTCAACGACACCGCCAAACCGGTGAGATCACTGTCCAAAGTGGCCGGAATTAT 216
QY 211 CCGGTGGTATGCTCTCCATGCTTACCTTCTCAACAACCTCTTATTCAGCTTATG 270
Db 217 CCCGCTGTTTATTTCTCCATGCTTATTTATCTTCCCACTACTCTGACGTTCTT 276
QY 271 TTGCAATGCTCTTTCATAGCTTTCATCTCAATCGCTCTCTGTTATATAGATGCCGGA 330
Db 277 AACCAATCGCTCCGATGTTTACATCTTGTAGCCCAAGTTGTGCAAAATTAATGCGG 336
QY 331 CCAGACACAATGATGATTAATCAACGCGGAGATTATGATTTGTTATCAGTAGA 390
Db 337 CCGGAGGGCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 396
QY 391 CTTAATCACTTTCTTCAGCGCAAGTAAACCAAACTATCCAAATTTGCCCTCCGCG 450
Db 397 CTCAAGCTCACTCACTCACTTCGTTAATGCTATGAAATACACCTCACTCGTGGC 456
QY 451 CATGCGCGGCGGAGAAACCGGCTTGGCGGCTTAAAGAAATTTG---GGTACTCC 507
Db 457 CACAGCCCGGGGGGAAACCGGCTTGGCGGCTTAAAGCAATTTAGAC 516
QY 508 TCGAATCTAAAGATCTGACATTTGATCGGTATAGATCAATGATGAAACGAGAAAGG 567

Db 517 CCATCCATCAAGCTTTTACGCTTATATAGAAATGATCAAGTCCGAGAACTTAACAAATAC 576
QY 568 AACAACACCCCTCTCCGCGGTGGCTTACCTTCAAACTCATTTGAACAGCAAAACG 627
Db 577 ATTAGAACCATCCGATATCTTAACTATTAACCGGAATCTTTGAGCTGACATACCG 636
QY 628 CCTATCTTGTGATCGGTTGGGGCTTGTGAACCCCTCGGAACCATTAATTCGCCCG 687
Db 637 GTTGCAGTGTG---GGAACCGGCTCG---ACCGAAGTGAACCACTGATCCACCA 690
QY 688 TGTGCACTCCCGGAGTGAATCAACGAGATCTTTTGGGAATGTCAAGGTCCAGATGG 747
Db 691 TGGCCACCAACGACCTTAACCAATGAGAGATTTACAAAGTGTAAAGCGACCAAGCC 750
QY 748 CATTCGTTGGGAAGATTATGCGCATTTGACATGCTTGTATGATGATACAAAGGAT 807
Db 751 CATTCGTTGGCTGCGGATTAACGACATATGATGATGATGATGATGATGATGATGAT 810
QY 808 AGAGGGAAGATCTTATTTGTTGTGTAAGATGCTGA---AGAGAGAGACCAATGAG 864
Db 811 GTTGGGTTTATGCGCGGTGTGTATGTGAAGATGGCAAGAAAGAAAGTCTGAGATAG 870
QY 865 AGATTGCTTGGTGGAGCTTGTGTATCATTTTGAAGGCTTATTTGGAAGAGA 917
Db 871 AGCTTTGATGGTGAATGTGTGTTGCTTCTCAAGTATGTTGTGGGGTGA 923

RESULT 12

ADJ98180
ID ADJ98180 standard; DNA, 1188 BP.

ADJ98180;
AC

DT 06-MAY-2004 (first entry)

DE Thale cress chlorophyllase 1 DNA.

XX phyto kinase; tocopherol biosynthesis; plant; drought resistance;

XX Thale cress; ds; chlorophyllase 1.

XX Arabidopsis thaliana.

PN WO2004013312-A2.

PR 05-AUG-2003; 2003WO-US025276.

PR 05-AUG-2002; 2002US-0400689P.

PR 05-AUG-2003; 2003US-00634548.

PA (MONS) MONSANTO TECHNOLOGY LLC.

PI Norris SR, Lincoln K, Abad MS, Billers R, Hartsuyker KK;

PI Hirschberg J, Karunanandaa B, Moshiri F, Stein JC, Valentin HE;

PI Venkatesh TV;

DR WPI; 2004-157125/15.

PT New phyto kinase polynucleotides, useful in mediating tocopherol

PS biosynthesis and in producing plants with increased drought resistance.

XX Claim 23; SEQ ID NO 18; 189pp; English.

XX The invention relates to a novel substantially purified nucleic acid

XX molecule encoding a phyto kinase. The nucleic acid molecules and

XX polypeptides of the invention may be useful in mediating tocopherol

XX biosynthesis and in producing plants with increased drought resistance.

XX The current sequence is that of the thale cress chlorophyllase 1 DNA of

XX the invention.

XX Sequence 1188 BP; 327 A; 255 C; 281 G; 325 T; 0 U; 0 Other;


```

Db      511 GCCCATGTAATCTTGCGAAACATTTCTCTCATTTCTCACTGGCATGTGCCAATCTTCA 570
Qy      614 ACTTAGACAAAACGCTTACTTGTGATCGGTGCGGGCTTGTAACCGCTCGAAC 673
Db      571 ATTGAAACATA---CCCATTTGTTAATTGGCACTGGGCTTAGGCCAGAGAAGCTAATT 627
Qy      674 CATTTATCCACCGTGTGACCTCCCGAGTGAATCACCGAGTCTTTTGGGAAATGTC 733
Db      628 TTTTATTCACCACTGTCTCTCTGATGGGTAACCAATAGAGATTTTCAATTAATGCA 687
Qy      734 AAGGTCCAGCATTTGCTTGTGGAAGATTAATGCGCATTTGGACATGCTTGATGATG 793
Db      688 AACCCCTTGTGCACATTTTGTGCACTGAGTATGTCATGACATGATGATGATG 747
Qy      794 ATCAAAAAGGATTAAGAGG---AAGAGTCTTATTTGTTGTGAAGATGCTGAAG--- 847
Db      748 TGACACCTGGCTTAATTGGGTCAATATTGTCAAAATTGTATATGCAAGAGTGGAAAGGCTC 807
Qy      848 AGAGAGACCAATGAGAGATTTGTTGTGACCTTGTGTATCATTTTGAAGGCTTATT 907
Db      808 CTAGGACTTGAATGAAAGAACCTGAGAGGGTGTGTGCTTTTAAAGGCACAGT 867
Qy      908 TGAAGG 914
Db      868 TGAATGG 874

```

RESULT 14

AAD38082
ID AAD38082 standard; cDNA; 1242 BP.

XX AC AAD38082;

DT 10-SEP-2002 (first entry)

XX Wheat chlorophyllase cDNA from wle1n.pk0058.a4:file clone.

XX Wheat; chlorophyllase; chlorophyll degradation; plant cell senescence;

KM enzyme; gene; ss.

XX Triticum aestivum.

XX Key Location/Qualifiers

FT CDS 47..1106

FT /tag= a
FT /product= "Wheat chlorophyllase protein from
FT wle1n.pk0058.a4:file clone"
FT /EC_number= "3.1.1.14"

XX WO200229022-A2.

XX 11-APR-2002.

XX 04-OCT-2001; 2001WO-US031059.

XX 05-OCT-2000; 2000US-0238161P.

XX (DUPO) DU PONT DE NEMOURS & CO E I.

XX Cahoon EB, Cahoon RE, Thorpe CJ;

XX WPT: 2002-444102/47.

XX P-PSDB; AAE23784.

XX An isolated polynucleotide encoding a plant chlorophyllase polypeptide,
XX used to produce transformed plants that have controlled induction or
XX postponement of senescence.

XX Claim 5; Page 59; 69pp; English.

XX The invention relates to plant chlorophyllase (EC 3.1.1.14) proteins and
XX polynucleotides encoding such proteins. Chlorophyllase enzyme is involved

CC in chlorophyll degradation in plant cell senescence. Sequences of the
CC invention are used to produce a plant by transforming a plant cell with
CC chlorophyllase and regenerating a plant from the transformed plant cell.
CC They may also be used to transform cells. The plants that are produced
CC can have economical importance as they can allow for the controlled
CC induction or postponement of senescence. The present sequence is wheat
CC chlorophyllase cDNA from wle1n.pk0058.a4:file clone

XX Sequence 1242 BP; 266 A; 382 C; 343 G; 251 T; 0 U; 0 Other;

Query Match 13.1%; Score 148.2; DB 6; Length 1242;

Best Local Similarity 55.2%; Pred. No. 9.7e-28;

Matches 357; Conservative 0; Mismatches 278; Indels 12; Gaps 3;

```

Qy      151 GCTTACCGTCTCCGCCAAGACGCTTGTGTGCTACGCGGTGGAGGAAGATTAAT 210
Db      173 GCGGCGCGGACACCAACGATCCCGGTCTGATCTGACACCAAGATGCGAAGCTAC 232
Qy      211 CCGGTGGATAGCTCTCCATGTTACTTCTTCAACTCTTCTATTTCTCAGCTTATG 270
Db      233 CCGGTGGATAGCTCTTGTGACGCTTCTTCTCTCAATACCACTTCTACGAACCTTCTC 292
Qy      271 TTGCATGTCTTCTCATAGGCTTCATCTCATGCTCTCTCAAT---ATATATGATGCGC 327
Db      293 CGGCACGTGGATCCACGCGCTTCATCTATGTCGCGCCCGCAGTTCAGCATCATATATA 352
Qy      328 GGACCAACACATGATGATGATTAATCAACGCGCGAGATTAATGATGTTATCATGA 387
Db      353 CTTTGGGTGACGCGAGGACATGCGCGGACCAAGGTGGAGACTGGCTCCCGGAC 412
Qy      388 GGACTTAATCACTTCTTCCAGCGCAATGAACACCAACCTATGCAATTTGCGCTTCC 447
Db      413 GGCCTCCCGTCCGTGCTGCCAAGGGGTGAGCGGAGCTCTGGAAGCTGCTTGCGCC 472
Qy      448 GGCATAGCCGCGGTGCAAAACCGGTTGCGGTGCTTAAAGAAATTTGGTACTCC 507
Db      473 GGCACAGCCGAGGAGGCCACACGCGCTTCTCCCTGAGCTTGGGGCACGCCAAGACC-- 529
Qy      508 TCGAATCTAAAGATCTGACATTAATGATGATTAATGATCAGTCATGGAACGGAAGG 567
Db      530 ---CAGCTAACCTTCTCCGCGCTCATGCACTGACCCCGCGGACGCGGAAGTCC 586
Qy      568 AAACAAACCCCTCTCCGCTGTTGGCTTAACCTTCAACTATTTGAACCTAGACAAAACG 627
Db      587 TCCAGCTCCAGCCCAAGATCTCACTTAGAGCGCTCTCTCCGCGATG---CGATG 643
Qy      628 CTTATCTTGTGATCGGTTGGGGCTTGTGAACCGCTCGAAACCAATTAATTTCCACCG 687
Db      644 CCGGTGCTGTCATCGGACCGGGCTCGGCGAGGAAGAAAGAAATATTTCTCCCTCC 703
Qy      688 TTGTCACCTCCCGAGATCAACCGAGATTTCTTTGGGGAATGTCAAGTCCAGCATGG 747
Db      704 TGCACACCAAGAGATGAACACGCGAGTTCTTACCGCGAGTCCAGCGCGCTGTAC 763
Qy      748 CATTTGTTGGAAGATTAATGAGCATTTGGACATGCTTGATGATGA 794
Db      764 TACTTTGACCAAGACTAGCGGCATCTGACATGCTGACAGACA 810

```

RESULT 15

AAD38078
ID AAD38078 standard; cDNA; 1104 BP.

XX AC AAD38078;

DT 10-SEP-2002 (first entry)

XX Soybean chlorophyllase cDNA from sfl1n1.pk002.m10:file clone.

XX Soybean; chlorophyllase; chlorophyll degradation; plant cell senescence;

KM enzyme; gene; ss.

XX Glycine max.

XX Key Location/Qualifiers
 FH 26..967
 FT /tag= a
 FT /product= "Soybean chlorophyllase protein from
 sfln1.pk002.m10:file clone"
 FT /EC_number= "3.1.1.14"
 XX
 PN MO200229022-A2.
 XX
 PD 11-APR-2002.
 XX
 PF 04-OCT-2001; 2001WO-US031059.
 XX
 PR 05-OCT-2000; 2000US-0238161P.
 XX
 PA (DUPO) DU PONT DE NEMOURS & CO E. I.
 XX
 PI Cahoon EB, Cahoon RE, Thorpe CJ;
 XX
 DR MPI: 2002-444102/47.
 DR P-PSDE; AAE23780.
 XX
 PT An isolated polynucleotide encoding a plant chlorophyllase polypeptide,
 PT used to produce transformed plants that have controlled induction or
 PT postponement of senescence.
 XX
 PS Claim 5, Page 53; 69pp; English.
 XX
 CC The invention relates to plant chlorophyllase (EC 3.1.1.14) proteins and
 CC polynucleotides encoding such proteins. Chlorophyllase enzyme is involved
 CC in chlorophyll degradation in plant cell senescence. Sequences of the
 CC invention are used to produce a plant by transforming a plant cell with
 CC chlorophyllase and regenerating a plant from the transformed plant cell.
 CC They may also be used to transform cells. The plants that are produced
 CC can have economical importance as they can allow for the controlled
 CC induction or postponement of senescence. The present sequence is soybean
 CC chlorophyllase cDNA from sfln1.pk002.m10:file clone
 XX
 SQ Sequence 1104 BP; 308 A; 225 C; 249 G; 322 T; 0 U; 0 Other;
 Query Match 12.8%; Score 145.4; DB 6; Length 1104;
 Best Local Similarity 53.8%; Pred. No. 4.9e-27;
 Matches 421; Conservative 0; Mismatches 341; Indels 21; Gaps 5;
 QY 143 CTCTAAGCTTCAACCGTCCGCAAGAGAGCTGTGGCTACGCGCGTGGAGAG 202
 DB 108 CATCAGTGTCTCTCTCACTCCAAACCATGTGTAATTTTACACCAACTGTCCTG 167
 QY 203 GAATTTATCGGAGTGTATGTCTCTCAAGTGAATCTTCTGTAACCTCTTCTATTC 262
 DB 168 GCGATACCTGTAAATTTGTCTCAAGCTTTTCAATTCGCAATTTCTACTACTCAA 227
 QY 263 AGCTTATGTGATGTCTCTCTCATGAGCTTCATCTCATGCGCTCCAGTTATAGTA 322
 DB 228 AGCTCCCTAGCCCATATGTCTCAGATGATTCATATGCTTGTCTCAACTGTTTCCA 287
 QY 323 TCGCCGAGAC--AGACACAATGATGATGAATTAATCAAGCGAGATTATGATTGT 379
 DB 288 ATGGGCTTCTATGTATGAGACCCACTGAAGTGAATATGACGAAAGTTGGGATTGA 347
 QY 380 TATCAGTAGAATTAATCACTTTTTCAGGCGCAAGTAACCAACCTATCCAAATTGG 439
 DB 348 TAGCTGAGAGCTTCAACATTTGCTTCCAGGAAGCTTGAAGCCAAATTTGACCAACTGG 407
 QY 440 CCTCTCCGCGCATAGCGCGGTGGCAAAACCGGTTTGGCGGTCCCTTAAGAAATTGG 499
 DB 408 TTCTATAGGTCAAGTAGAGGGGTGGAAACTGTATTTGCTGTGCTCT-----TGCTC 461
 QY 500 GGTACTCTCTGAATCTAAAGATCTCGACATTTGATCGGTATAGATCAAGTGAAGACAG 559
 DB 462 ATGCAAAACCTAATCTCAAGTTTTCAGACCTTGTAGGCATAGACCTGTGGCTGGACAT 521

QY 560 GGAAGGGAACAAACCCCTCTCCGAGTGTGGCTTACCTCCAACTCATTTGACTAG 619
 DB 522 CTAAATATTTAGAACAGTCTCTCATATTTCTACCTGGCAAGCCACGGTCTTGATTT-- 579
 QY 620 ACAAAAGCCTATPACTTGTATCGATTGCGGAGCTTGGTGAACCGCTCGGAACCATTA 679
 DB 580 -GAAATGCCAGTTGAAGTAAATTTGGCACTGATTTGGCCCAAGAAAGCTTAATTTGTTGA 638
 QY 680 TCCACCGTGTGACCTCCGAGTGAATACCGAGATTCTTTCGGAAATGTCAAGTCT 739
 DB 639 CTCACCGTGTGCTCTGATGGGGTGAATATAAGAGTTCTTCAACGAGTGGAAAACCC 698
 QY 740 CAGATGGCATTTCTGTGCAAGATTAAGGCAATTTGGACATGCTTGAATGATGATCA 799
 DB 699 CTGTGCTAAATTTGTTGTAGCAAAAGTATGTCACATGAGCATGTTGAATGATGACAC 758
 QY 800 AAGGATTTAGAGGA--AGAGTTCTATTTGTTGTAGAAATG-----TGAAGAGA 850
 DB 759 CAGGCTTAATTTGGACATTTGTGTCAAAAGTATGTATGTATGAATGGACAGCGGTCTTA 818
 QY 851 GGAACCAATAGAGAGATTTGTTGTGAGCTTGTGTATCATTTTGAAGGCTTATTTGG 910
 DB 819 GGAATTGATGAGAGAGACACTGAGAGGTTGTTGTGCTCTTGTGAGGGCAATTTGA 878
 QY 911 AAG 913
 DB 879 ATG 881

Search completed: March 20, 2006, 14:14:24
 Job time : 671.906 secs

THIS PAGE LEFT BLANK

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocelebration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 20, 2006, 14:06:13 ; Search time 1051.94 Seconds
(without alignments)
8922.325 Million cell updates/sec

Title: US-10-634-548-19

Perfect score: 1135
Sequence: 1 aaaaaaagtaagaagaagaa.....aaaaaaaaaaaaaaaa 1135

Scoring table:
IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Published Applications NA Main:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1135	100.0	1135	US-10-634-548-19	Sequence 19, Appl
2	367.4	32.4	1216	US-10-381-123-3	Sequence 3, Appl
3	256.8	22.6	1444	US-10-381-123-13	Sequence 13, Appl
4	237.4	20.9	987	US-10-381-123-9	Sequence 9, Appl
5	201.6	17.8	1302	US-10-381-123-5	Sequence 5, Appl
6	201.6	17.8	2439	US-10-425-115-140359	Sequence 140359, A
7	185.6	16.4	796	US-10-424-599-86252	Sequence 86252, A
8	184	16.2	1125	US-10-425-114-26342	Sequence 26342, A
9	179.4	15.8	635	US-10-424-599-131410	Sequence 131410, A
10	163.4	14.4	975	US-09-938-842A-158	Sequence 158, App
11	163.4	14.4	975	US-09-938-842A-158	Sequence 158, App
12	163.4	14.4	1188	US-10-634-548-18	Sequence 18, Appl
13	151	13.3	1174	US-10-381-123-7	Sequence 7, Appl
14	148.2	13.1	1242	US-10-381-123-17	Sequence 17, Appl
15	145.4	12.8	1104	US-10-381-123-9	Sequence 9, Appl
16	142.8	12.6	1125	US-10-381-123-11	Sequence 11, Appl
17	135.4	11.9	1307	US-10-424-599-45715	Sequence 45715, A
18	128.2	11.3	1155	US-10-437-963-94232	Sequence 94232, A
19	116.2	10.2	751	US-10-425-115-70835	Sequence 70835, A
20	116	10.2	1244	US-10-425-114-25386	Sequence 25386, A
21	91.4	8.1	1068	US-10-425-114-25387	Sequence 25387, A
22	88.4	7.6	1156	US-10-424-599-34940	Sequence 34940, A
23	88.4	7.6	1156	US-10-381-123-25	Sequence 25, Appl

24	85.2	7.5	538	7	US-10-424-599-34941	Sequence 34941, A
25	80.8	7.1	512	8	US-10-425-115-41107	Sequence 41107, A
26	71.4	6.3	916	9	US-10-381-123-15	Sequence 15, Appl
27	68.8	6.1	1004	7	US-10-424-599-68342	Sequence 68342, A
28	67.2	5.9	708	7	US-10-437-963-34072	Sequence 34072, A
29	52.2	4.6	1236	8	US-10-425-115-166593	Sequence 166593, A
30	52.2	4.6	1274	9	US-10-381-123-27	Sequence 27, Appl
31	49	4.3	516	7	US-10-021-323-6457	Sequence 6457, Ap
32	47.4	4.2	487	7	US-10-021-323-2841	Sequence 2841, Ap
33	47.4	4.2	467	7	US-10-021-323-6541	Sequence 6541, Ap
34	47.4	4.2	552	7	US-10-021-323-93	Sequence 93, Appl
35	46.8	4.1	579	8	US-10-357-930-23827	Sequence 23827, A
36	46.8	4.1	579	8	US-10-357-930-29732	Sequence 29732, A
37	46.4	4.1	10039	6	US-10-311-455-2016	Sequence 2016, Ap
38	45.8	4.0	419	3	US-09-960-352-11224	Sequence 11234, A
39	45.8	4.0	629	6	US-10-125-968-78	Sequence 78, Appl
40	45.8	4.0	3673778	6	US-10-112-841-1	Sequence 1, Appl
41	45.6	4.0	728	7	US-10-767-701-3821	Sequence 3821, Ap
42	45.4	4.0	391	8	US-10-357-930-8140	Sequence 8140, Ap
43	45.4	4.0	524	7	US-10-767-701-3830	Sequence 3830, Ap
44	45.4	4.0	626	8	US-10-357-930-60960	Sequence 60960, A
45	45.2	4.0	421	3	US-09-918-995-7003	Sequence 7003, Ap

ALIGNMENTS

RESULT 1
US-10-634-548-19
Sequence 19, Application US/10634548
Publication No. US20040045051A1
GENERAL INFORMATION: US20040045051A1r1s, Susan R
APPLICANT: No. US20040045051A1r1s, Susan R
APPLICANT: Lincoln, Kim
APPLICANT: Abad, Mark Scott
APPLICANT: Eilers, Robert
APPLICANT: Hartsyker, Karen Kindle
APPLICANT: Hirsberg, Joseph
APPLICANT: Karunanandaa, Balasubjini
APPLICANT: Moshari, Farhad
APPLICANT: Stein, Joshua C.
APPLICANT: Valentin, Henry E.
APPLICANT: Venkatesh, Tyamagondlu V.
TITLE OR INVENTION: Tocopherol biosynthesis related genes and used thereof
FILE REFERENCE: Ren-01-125
CURRENT APPLICATION NUMBER: US/10/634,548
CURRENT FILING DATE: 2003-08-05
PRIOR APPLICATION NUMBER: us 60/400,689
PRIOR FILING DATE: 2002-08-05
NUMBER OF SEQ ID NOS: 79
SOFTWARE: PatentIn version 3.2
SEQ ID NO 19
LENGTH: 1135
TYPE: DNA
ORGANISM: Arabidopsis thaliana
US-10-634-548-19

Query Match 100.0%; Score 1135; DB 7; Length 1135;

Best Local Similarity 100.0%; Pred. No. 8.6e-299;

Matches 1135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	AAAAAAGTAAAGAAAAAAGAAAAAAGAAAAAAGTCTCTTCTTCATCA	60
DB	1	AAAAAAGTAAAGAAAAAAGAAAAAAGAAAAAAGTCTCTTCTTCATCA	60
QY	61	AGAAAGCCCTTGAAGTGAATCAATCAATCTTGAAGTGAATCATCT	120
DB	61	AGAAAGCCCTTGAAGTGAATCAATCAATCTTGAAGTGAATCATCT	120
QY	121	CGTTGCTGCAAAATTAACCGTCTTGAAGTTCACCGTCTCGCAAGCAGCTGTG	180
DB	121	CGTTGCTGCAAAATTAACCGTCTTGAAGTTCACCGTCTCGCAAGCAGCTGTG	180

```

QY 181 GTGGCTAGCCGGTGGAGGAGAGATTATCCGGGTGATGCTCCATGTTACCTT 240
Db 181 GTGGCTAGCCGGTGGAGGAGAGATTATCCGGGTGATGCTCCATGTTACCTT 240
QY 241 CTCTCAACATCTCTTATCTCAGCTTATGTTGACATGCTCTTCTCATGCTTCACTCC 300
Db 241 CTCTCAACATCTCTTATCTCAGCTTATGTTGACATGCTCTTCTCATGCTTCACTCC 300
QY 301 ATCGCTCTCAGTATATATGATATGCGCGGACCAACACAAATGATGATTAATCAACG 360
Db 301 ATCGCTCTCAGTATATATGATATGCGCGGACCAACACAAATGATGATTAATCAACG 360
QY 361 GCGGAGATTATGATGTTGTTATCAGTAGGACTTATATCACTTCTCCAGCGCAATTAACA 420
Db 361 GCGGAGATTATGATGTTGTTATCAGTAGGACTTATATCACTTCTCCAGCGCAATTAACA 420
QY 421 CCAAACTATCCAAATTTGCCCTCTCCGGCCATAGCCCGGCGTGGCAAAACCGCGTTTGC 480
Db 421 CCAAACTATCCAAATTTGCCCTCTCCGGCCATAGCCCGGCGTGGCAAAACCGCGTTTGC 480
QY 481 GTCCGCTTAAAGAAATTTGGGTACTCTCGAATCTAAAGATCTCGAATGATCGTATTA 540
Db 481 GTCCGCTTAAAGAAATTTGGGTACTCTCGAATCTAAAGATCTCGAATGATCGTATTA 540
QY 541 GATCCAGTGTGATGAAACAGGAGAAAGGAAACAAACCCCTCCGCGTGTGGCTTACCTT 600
Db 541 GATCCAGTGTGATGAAACAGGAGAAAGGAAACAAACCCCTCCGCGTGTGGCTTACCTT 600
QY 601 CCAAACTCATTTGACCTTAGACAAACAGCCTATCTGATGATGATGATGATGATGATGAT 660
Db 601 CCAAACTCATTTGACCTTAGACAAACAGCCTATCTGATGATGATGATGATGATGATGAT 660
QY 661 ACCGCTCGGAACCCATATTTCCACCGGTGACACTCCCGGATGATATCAACGAGATTC 720
Db 661 ACCGCTCGGAACCCATATTTCCACCGGTGACACTCCCGGATGATATCAACGAGATTC 720
QY 721 TTTCCGGAATGTCAAGTCTCAGCATGCTTCTGTTGCGAAGATTTATGCGCATTTGAGC 780
Db 721 TTTCCGGAATGTCAAGTCTCAGCATGCTTCTGTTGCGAAGATTTATGCGCATTTGAGC 780
QY 781 ATGCTTGTATGATATCAAAAGGAAATTAAGGGAAGATTTTATGTTTGTGTAAGAT 840
Db 781 ATGCTTGTATGATATCAAAAGGAAATTAAGGGAAGATTTTATGTTTGTGTAAGAT 840
QY 841 GGTGAAGAGAGAGACCAATGAGAGATTCGTTGCGATCTTGTGATCAATTTTGAAG 900
Db 841 GGTGAAGAGAGAGACCAATGAGAGATTCGTTGCGATCTTGTGATCAATTTTGAAG 900
QY 901 GCTTATTTGGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
Db 901 GCTTATTTGGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
QY 961 GTTCCGTTGAATTCAGAGATTTGAGGTTATCATGTAACATAGTTTCTTAAAGGG 1020
Db 961 GTTCCGTTGAATTCAGAGATTTGAGGTTATCATGTAACATAGTTTCTTAAAGGG 1020
QY 1021 CTGCTTATTTCTATGTCATATCATCATGCTTGTGCTTATGCTTATGCTTATGCTTAT 1080
Db 1021 CTGCTTATTTCTATGTCATATCATCATGCTTGTGCTTATGCTTATGCTTATGCTTAT 1080
QY 1081 TGTACAACTCTTAAAGTCACTCTTGTGCTTACAAAATTAATTAATTAATTAATTAAT 1135
Db 1081 TGTACAACTCTTAAAGTCACTCTTGTGCTTACAAAATTAATTAATTAATTAATTAAT 1135

```

RESULT 2
US-10-381-123-3

; Sequence 3, Application US//10381123

; Publication No. US20050081263A1

; GENERAL INFORMATION:

; APPLICANT: E.I. du Pont de Nemours and Company

; TITLE OF INVENTION: Chlorophyllase

; FILE REFERENCE: BBI477 PCT

```

; CURRENT APPLICATION NUMBER: US/10/381,123
; ;
; ; CURRENT FILING DATE: 2003-03-21
; ; PRIOR APPLICATION NUMBER: PCT/US01/31059
; ; PRIOR FILING DATE: 2001-10-04
; ; PRIOR APPLICATION NUMBER: 60/238,161
; ; PRIOR FILING DATE: 2000-10-05
; ; NUMBER OF SEQ ID NOS: 30
; ; SOFTWARE: Microsoft Office 97
; ; SEQ ID NO 3
; ; LENGTH: 1216
; ; TYPE: DNA
; ; ORGANISM: Vitis sp.
; ; US-10-381-123-3
;
Query Match 32.4%; Score 367.4; DB 9; Length 1216;
Best Local Similarity 63.9%; Pred. No. 3.5e-89;
Matches 608; Conservative 0; Mismatches 331; Indels 12; Gaps 3;
;
QY 50 CTCTCTATCAAGAAAGCCCTTGAAGATGAGCAAAATCAATCAATCTTAACTTTG 109
Db 1 CTCTCAATGTTCCACACATGCTGCAAAATGTTTGAATTTGAAAGACATCAAGTGC 60
QY 110 ACTCATCATCTCTGTCGCAAAATTAACACCGTCTTCAAGCTTCAACGCTCCGCCAA 169
Db 61 TTCTAAGAGTGAACAGGACCTTGACACCAAGTCAATCTTCTGTTCCCTCCAC 120
QY 170 AGCAGCTGTTGGTGTACACCGCGTGAAGAGATTAATCCGCTGATGCTCTCC 229
Db 121 TGCAACTCTGATGTTCAAGCCCTCAAGACAGGAGATTCGCGCTGCTCTCTTC 180
QY 230 ATGCTTACCTTCTTCAACATCTCTTATCTCAGCTTATGTTGATGATGATGATGATGAT 289
Db 181 ATGCTTATCTTCTTAAATCTTCTTCTCAGCTTCAATCAACATGACCTTCAATG 240
QY 290 GCTTCATCTCATGCTCTCAATTAATATATATGATGATGATGATGATGATGATGATGAT 349
Db 241 GTTTCATGTTCTTCTCTCAATTAATATATGATGATGATGATGATGATGATGATGATGAT 300
QY 350 TTAATCAAGCGGAGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 409
Db 301 TCAAGTCCGAGCTGCTTAAACAAATGTTATCAAAAGATCAATGATCTTCTCTC 360
QY 410 CGCAAGTAAACCAACCTATCCAAATTTGCGCTCTCCGCGATAGCCGCGTGGCAAAA 469
Db 361 CCATGTTCCGCAAAATTTAAGCAAACTAAGACTTGCCTGATGATGATGATGATGATGAT 420
QY 470 CCGGTTTGGCGTGCCTTAAAGAAATTTGGGATCTCTGAAATCTAAAGATCTGACAT 529
Db 421 CTGCTTATGCTCTGACATGAGAAA-----GCATCACTTCTGAAATTTCAAGCT 474
QY 530 TGAATCGTATTAATCCAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 589
Db 475 TGAATGCAATGAACCGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 534
QY 590 TGGCTTACCTTCAAACTCAATTTGACCTAGCAAAAGCCTATCTTGTGATCGGTTCCG 649
Db 535 TCACATGTTCTCTCATCTATTTGATCTAAGC---ATGCAATGATGATGATGATGATGATGAT 591
QY 650 GCGTGTGTAACCGCTCGGAACCCATTAATTTCCACCGTGTGACCTCCGGAATGATC 709
Db 592 GTTGGGTGAAGTAAAGAAACCTCTGTTCCCTCTGTCGCCCCCAAGGCGCTTAAAC 651
QY 710 ACCGAGTCTTCTCGGAATGTCAAAGTCCAGATGATGATGATGATGATGATGATGATGAT 769
Db 652 ATGAGACTTCTTAAAGAAATGCGTGAACCAAGCTTGTATTTCTTGCAGGAATATG 711
QY 770 GGCATTTGACATGCTTGAATGATGATCAAAAGGATTTAAGGAGAGATTTATTTGTT 829
Db 712 GCATCTTGAATCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 771
QY 830 TGTGTAAGATGTTGAAGAGAGACCAATGAGAGATTTGTTGTTGAGCTTGTGTAT 889
Db 772 TGTGTAAGATGTTGAAGAGAGACCAATGAGAGATTTGTTGTTGAGCTTGTGTAT 831

```


Oy 143 CTTCAAGAGCTTCAACCGTCTCGGCAAAACAGCTGTGGTGCCTACGCGGTGGAGGAG 202
 Db 241 CGTCAGCGCGGCGCAGCGTCCCGCCGAGCCGCTGCTCGCGCGCGCGCGGAGACTG 300
 Oy 203 GAGATTATCCGGTGGATGCTCTCTCCATGTTAACTTCTCTCAAACTCTCTTATTTCTC 262
 Db 301 GGGATATACCGGGATGATCTGTCTCTCAACAGGGCTACTCGCGGTCAACTCTCTTCACTCCC 360
 Oy 263 AGCTTATGTTGATGCTCTTCTCTCAAGGCTTCACTCTCACTGCTCTCTCAGTAAATAGTA 322

Query Match	17.8%;	Score 201.6;	DB 8;	Length 2439;
Best Local Similarity	54.6%;	Pred. No. 9.6e-44;		

Matches 449; Conservative 0; Mismatches 364; Indels 9; Gaps 2;

QY 143 CTTTAAAGCTTACCGCTTCGCCAAAGCACTGTGTGGCTTACCGCGGTGAGAG 202
 Db CGTCACGCGCGGAGCGCTTCGCCAAAGCGCTGTGTGTGCGCGCGCGGAGACTG 341

QY 203 GAGATTATCCGGTGTGATGCTCTCCATGCTTACCTTCTTAACTCTCTTATTTCTC 262
 Db GGGAGTACCGGGTATCTGTCTTACCGGCTTACCGGCTTACCTCTTCTTACTCC 401

QY 263 AGCTTATGTCATGTCTCTTCATGAGCTTCACTCATGCTCTCATGTTATATAGTA 322
 Db AGCTTGTGAGCAAGTGGCTTCCATGGCTTATCTGTGTGAGACTCAGCTGACACA 461

QY 323 TCGCCGACCAAGACAAATGATTAATCAACGCGCGAGATTATGATTTGTTAT 382
 Db TATCTGGGCGCGACCAACGAGAGATCACTCAGCGCGCGCTCATGCACTGGCTAG 521

QY 383 CAGTAGGACTTAACTTCTTCCAGCGCAATTAACCAACTTATCCAAATTTGCC 442
 Db CCACCGGAGCTGCGTCAACTCTGCACTCGCGCTCGCGCAACCTTAAACAGGTGCA 581

QY 443 TCTCCGCGCAATGCGCGGTGGCAAAACCGGCTTGGGTGCGCTTAAAGAAATTTGGT 502
 Db TCTCCGCGCAAGTCCGCGCGGAAAGGTGGGCTTGGCGCTGGCTTGGCGCAAG 641

QY 503 ACTCTCGAATCTTAAAGATCTGACATTTGATGATGATGATGATGATGATGATG 562
 Db CCAAGCTCGCTGTCTCTCGCGCGCGCTGCTGCGCGGAGACCGGTGAGACCGGATGGCG 701

QY 563 AAGGGAACAACCCCTCTCCGCTGTGTGCTTACTTCCAAATCTTATGACTTAACA 622
 Db TGCGCAAGCAACACCGCGCGCTGCTGCAAGGCAAGCGCTGCTGCACTGCGTGA 758

QY 623 AAACGCTTATCTTGTATCGGCTTGGGAGTGGTGAACCGCTCGGAAACCATTTATCC 682
 Db GTGCGCGCGCATGCTGATCGGCAAGGCGCTGCGCGAGCTGCGCGCGCTGCTGCTC 818

QY 683 CACCTGTGCACTCCCGAGTGAATCAACGAGGTTCTTTGGGAAATGTCAGG---- 737
 Db CGCGTCCGCGCGCGCGCGCTGATCAAGCGCGCTTCTTACGAGACTGCAAGCGCGCG 878

QY 738 -TCCAGATGGCATTTGTTGCGAAGATTATGGGCACTTGGACATCTGATGATATA 796
 Db CACCAAGGTGCACTGCTGTGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATG 938

QY 797 CAAAAGGATTAGAGGAGAGTCTTATTTGTTGTTAGAAATGTGAAGAGAGAC 856
 Db CGCGCGCGCGCGAGGAGATGCTCAACGCGCACTGCAAGAGGCGCGCGCGCGCGCG 998

QY 857 CAATGAGAGATTCGTTGTGAGACTTGTGATCAATTTTGAAGCTTATTTGAGAGAG 916
 Db CCATGCGCGCGCTTGTGCGCGCGCGCGCGCGCTGCTTCAAGAAATGAGTGGCTGGG 1058

QY 917 ATGATCGTGAATTAATTAATCAAGATGAGGTGTCAGAG 958
 Db 1059 ACGCGCGCGAGTGAAGATCAAGCGCGCGCGCGCGAGCAGG 1100

RESULT 7
 US-10-424-599-86252
 ; Sequence 86252, Application US/10424599
 ; Publication No. US20040031072A1
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa Thomas J
 ; APPLICANT: Kovalic David K
 ; APPLICANT: Zhou Yihua
 ; APPLICANT: Cao Yongwei
 ; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
 ; FILE REFERENCE: 38-21(53223)B
 ; CURRENT APPLICATION NUMBER: US/10/424,599
 ; CURRENT FILING DATE: 2003-04-28

NUMBER OF SEQ ID NOS: 285684
 SEQ ID NO 86252
 ; LENGTH: 796
 ; TYPE: DNA
 ; ORGANISM: Glycine max
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: PAT_MRT3847_48899C.1
 US-10-424-599-86252

Query Match 16.4%; Score 185.6; DB 7; Length 796;
 Best Local Similarity 65.5%; Pred. No. 1.2e-39;
 Matches 288; Conservative 0; Mismatches 149; Indels 3; Gaps 1;

QY 509 CGAATCTAAAGATCTGACATTTGATGATGATGATGATGATGATGATGATGATG 568
 Db CTATCTCAAGTTTTCAGCTTATGAGTGTGATCAGTGTGATGATGATGATGATGATG 67

QY 569 AAGAAACCCCTCTCGCGCTTGGCTTACCTTCAACCTCATTTGACATGACAAACGC 628
 Db AGCAACCCCGCACCTGTTCTACCTATGTTCTTAACTCATTTGATTTTG--ATATGG 124

QY 629 CTATCTTGTGATCGGTTGCGGCTTGTGAAACCGCTCGAACCATTATTTCCACCGT 688
 Db 125 CAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 184

QY 689 GTGCACCTCCCGAGTGAATCAACGAGAGTTCTTTGGGAAATGCAAGTCCGACATGGC 748
 Db 185 GCGTCCGTAAAGGTGTCAACATGAAAACTTCTTCAATGATGATGATGATGATGATG 244

QY 749 ATTTGTTGGAAGATTTATGAGCATTTGACATGCTTATGATGATGATGATGATGATG 808
 Db 245 ATTTGTTGGAAGATTTATGAGCATTTGACATGCTTATGATGATGATGATGATGATG 304

QY 809 GAGGAAAGATTTCTTATTTGTTGTGATGATGATGATGATGATGATGATGATGATG 868
 Db 305 GGGGAAAGATTTCTTATTTGTTGTGATGATGATGATGATGATGATGATGATGATG 364

QY 869 TCGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 928
 Db 365 TTGTTGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 424

QY 929 TAGTTAAGATCAAGATGGG 948
 Db 425 TGTGACATTAAGAGACAG 444

RESULT 8
 US-10-425-114-26342
 ; Sequence 26342, Application US/10425114
 ; Publication No. US20040034888A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Liu, Jingdong
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Screen, Steven E.
 ; APPLICANT: Tabaska, Jack E
 ; APPLICANT: Cao, Yongwei
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 ; FILE REFERENCE: 38-21(5313)B
 ; CURRENT APPLICATION NUMBER: US/10/425,114
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 73128
 ; SEQ ID NO 26342
 ; LENGTH: 1125
 ; TYPE: DNA
 ; ORGANISM: Brassica napus
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: LIB4156-010-D3_FLI
 US-10-425-114-26342

Query Match 16.2%; Score 184; DB 7; Length 1125;
 Best Local Similarity 56.1%; Pred. No. 3.9e-39;


```

151 GCTTACCGTCTCCGCAAGAGAGCTTTGGTCTACCGCGTGGAGAAAGATTTAT 210
109 GATTCAACAGACCCGCAAAACCGGTGAGAACTCTGTCCAAAGTCCGGAATTTAT 168
211 CCGGTGTGATGCTCTCCATGCTTACCTTCTCAACATCTCTTATTTCTAGCTTATG 270
169 CCGGTGTTTTATTTCTTCATGCTTTTATCTTGCAACTACTTCTGACGCTTCTT 228
271 TTGATGCTCTCTTCTCATGCTTATCTCATGCTCTCTCATGTTATATGATGCGGGA 330
229 AACCAATCGCTTCCATGCTTATCACTCTGAGCCCAAGTTGCAAAATTTATGCGG 288
331 CCAGACAAATGATGATGATTAATCAACGCGGAGATTATGATGTTGATTCATGATAGA 390
289 CCGGAGGCGCAAGTGGAGTGGACGATCTGGAACTGTATTAATCTGGGCAATCGAAAC 348
391 CTATATCACTTCTTCCAGCGCAAGTAAACCAAACTATCCAAATTTGCCCTCCGGC 450
349 CTCAAGCTCACTACCAACTTCCGTAAATGCTAAATGGAATAACACTCTCACTCGTGGC 408
451 CATAGCCGCGGTGGCAAAACCGCGTTTGCCTTAAAGAAATTTG--GGTACTCC 507
409 CACAGCCGCGGTGGGAAAACGCGCTTTCGGTTGCGCTAGGCAATGCGCAACATTTAGAC 468
508 TCGAATCTAAAGATCTCGACATGATCGGTATTAATCCAGTCTGATGGAACAGGAAAGG 567
469 CATTCATCAAGTCTTCACTCTTAATGAAATGATTCAGTCCGAGAACTTCAACAAATAC 528
568 AAACAAACCCCTCTCCGCGTGTGGCTTACCTTCCAACTCATTTGACCTTACGAAAC 627
529 ATTGAACCGATCCGCAATCTTAAAGTAAACCGGAATCTTTCGAGCTGGAATACCG 588
628 CTTATCTTGTGATGCTTGGGCGTGTGTAACCGCTCGAAACCATTAATCCACCG 687
589 GTTGAGAGTGTG--GGAAACCGGACTCGG--ACGAAAGTGAACAACGATGATCCACA 642
688 TGTGACCTCCGCGAGTGAATCAACGAGATTTCTTCCGGAATGTCAGATCCAGCATG 747
643 TCGCACCAACGAGACTTAAACATGAGAGATTTTCAAAAGAGTGAAGCGACGAAAGCC 702
748 CATTTCGTTGCGAAGATTAATGAGCAATTTGACATGCTTGAATGATACAAAGAGATT 807
703 CATTTCGTTGCGGAGATTAATGAGCAATGATGATGATGATGATGATGATGATGATG 762
808 AGAGGGAAGAGTCTTATTTGTTTGTGTAAGATGTGA--AGAGAGAGACCAATGAG 864
763 GTTGCGTTTATGCGCGGTGTGTATGTGTAAGATGCGCAAAAGAAAGTCTGAGATGAG 822
865 AGATTGCTGTGAGACTTGTGTATCATTTTGAAGGCTTATTTGGAAGAGA 917
823 AGCTTTAGGTGAATGTGTGCTGCTTCTCAAGTATGATTTGTGGGGTGA 875

```

RESULT 11
 US-09-938-842A-158
 ; Sequence 158, Application US/09938842A
 ; Publication No. US2004000947649
 ; GENERAL INFORMATION:
 ; APPLICANT: Harper, Jeff
 ; APPLICANT: Kreps, Joel
 ; APPLICANT: Wang, Xun
 ; APPLICANT: Zhu, Tong
 ; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
 ; TITLE OF INVENTION: SAME, AND METHODS OF USE
 ; FILE REFERENCE: SRIPI300-3
 ; CURRENT APPLICATION NUMBER: US/09/938, 842A
 ; PRIOR FILING DATE: 2001-08-24
 ; PRIOR APPLICATION NUMBER: US 60/227, 866
 ; PRIOR FILING DATE: 2000-08-24
 ; PRIOR APPLICATION NUMBER: US 60/264, 647
 ; PRIOR FILING DATE: 2001-01-16
 ; PRIOR APPLICATION NUMBER: US 60/300, 111
 ; PRIOR FILING DATE: 2001-06-22

```

; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 158
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-158

Query Match      14.4%; Score 163.4; DB 3; Length 975;
Best Local Similarity 54.3%; Pred. No. 1.5e-33;
Matches 420; Conservative 0; Mismatches 341; Indels 12; Gaps 4;

151 GCTTACCGTCTCCGCAAGAGAGCTTTGGTCTACCGCGTGGAGAAAGATTTAT 210
109 GATTCAACAGACCCGCAAAACCGGTGAGAACTCTGTCCAAAGTCCGGAATTTAT 168
211 CCGGTGTGATGCTCTCCATGCTTACCTTCTCAACATCTCTTATTTCTAGCTTATG 270
169 CCGGTGTTTTATTTCTTCATGCTTTTATCTTGCAACTACTTCTGACGCTTCTT 228
271 TTGATGCTCTCTTCTCATGCTTATCTCATGCTCTCTCATGTTATATGATGCGGGA 330
229 AACCAATCGCTTCCATGCTTATCACTCTGAGCCCAAGTTGCAAAATTTATGCGG 288
331 CCAGACAAATGATGATGATTAATCAACGCGGAGATTATGATGTTGATTCATGATAGA 390
289 CCGGAGGCGCAAGTGGAGTGGACGATCTGGAACTGTATTAATCTGGGCAATCGAAAC 348
391 CTATATCACTTCTTCCAGCGCAAGTAAACCAAACTATCCAAATTTGCCCTCCGGC 450
349 CTCAAGCTCACTACCAACTTCCGTAAATGCTAAATGGAATAACACTCTCACTCGTGGC 408
451 CATAGCCGCGGTGGCAAAACCGCGTTTGCCTTAAAGAAATTTG--GGTACTCC 507
409 CACAGCCGCGGTGGGAAAACGCGCTTTCGGTTGCGCTAGGCAATGCGCAACATTTAGAC 468
508 TCGAATCTAAAGATCTCGACATGATCGGTATTAATCCAGTCTGATGGAACAGGAAAGG 567
469 CATTCATCAAGTCTTCACTCTTAATGAAATGATTCAGTCCGAGAACTTCAACAAATAC 528
568 AAACAAACCCCTCTCCGCGTGTGGCTTACCTTCCAACTCATTTGACCTTACGAAAC 627
529 ATTGAACCGATCCGCAATCTTAAAGTAAACCGGAATCTTTCGAGCTGGAATACCG 588
628 CTTATCTTGTGATGCTTGGGCGTGTGTAACCGCTCGAAACCATTAATCCACCG 687
589 GTTGAGAGTGTG--GGAAACCGGACTCGG--ACGAAAGTGAACAACGATGATCCACA 642
688 TGTGACCTCCGCGAGTGAATCAACGAGATTTCTTCCGGAATGTCAGATCCAGCATG 747
643 TCGCACCAACGAGACTTAAACATGAGAGATTTTCAAAAGAGTGAAGCGACGAAAGCC 702
748 CATTTCGTTGCGAAGATTAATGAGCAATTTGACATGCTTGAATGATACAAAGAGATT 807
703 CATTTCGTTGCGGAGATTAATGAGCAATGATGATGATGATGATGATGATGATGATG 762
808 AGAGGGAAGAGTCTTATTTGTTTGTGTAAGATGTGA--AGAGAGAGACCAATGAG 864
763 GTTGCGTTTATGCGCGGTGTGTATGTGTAAGATGCGCAAAAGAAAGTCTGAGATGAG 822
865 AGATTGCTGTGAGACTTGTGTATCATTTTGAAGGCTTATTTGGAAGAGA 917
823 AGCTTTAGGTGAATGTGTGCTGCTTCTCAAGTATGATTTGTGGGGTGA 875

```

RESULT 12
 US-10-634-548-18
 ; Sequence 18, Application US/10634548
 ; Publication No. US20040045051A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Lincoln, Kim
 ; APPLICANT: Abad, Mark Scott
 ; APPLICANT: Bilers, Robert

APPLICANT: Hartsuyker, Karen Kindle
APPLICANT: Hirschberg, Joseph
APPLICANT: Kannanandaa, Balasubramini
APPLICANT: Moshiri, Farhad
APPLICANT: Stein, Joshua C.
APPLICANT: Valentin, Henry E.
APPLICANT: Venkatesh, Thyamagondlu V.
TITLE OF INVENTION: Tocopherol biosynthesis related genes and used thereof
FILE REFERENCE: Ren-01-125
CURRENT APPLICATION NUMBER: US/10/634,548
PRIOR FILING DATE: 2003-08-05
PRIOR APPLICATION NUMBER: us 60/400,689
NUMBER OF SEQ ID NOS: 79
SOFTWARE: PatentIn version 3.2
SEQ ID NO: 18
LENGTH: 1188
TYPE: DNA
ORGANISM: Synechocystis PCC6803
US-10-634-548-18

Query Match 14.4%; Score 163.4; DB 7; Length 1188;
Best Local Similarity 54.3%; Pred. No. 1.7e-33;
Matches 420; Conservative 0; Mismatches 341; Indels 12; Gaps 4;

QY 151 GCTTACCGCTCCGCGCAAGACGCTGTGGTACGCGCGTGAGGAGGATTTAT 210
DB 157 GATTCAACAGCACCGCCCAAAACCGGTAGATCCTGTCCACAGTCCCGGAACTTAT 216
QY 211 CCGGTGTGATGCTCCTCCATGTTACCTTCTCAAACTCCTTCTATCTCAGCTTATG 270
DB 217 CCGGTGTGATGCTCCTCCATGTTACCTTCTCAAACTCCTTCTATCTCAGCTTATG 276
QY 271 TTGCATGCTCTTCTCATGAGCTTCACTCATGCTCTCTCACTTATATAGTATGCGCGGA 330
DB 277 AACCATGCTCTTCCATGAGTTACATTTCTTGAAGCCCAAGTTGTGCAATTAATGCGG 336
QY 331 CCAGACACATGATGATGATTAATCAACGCGGAGATTAATGATTTGTTATCAGTAGGA 390
DB 337 CCGGAGGAGGCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 396
QY 391 CTTAATCACTTTCTTCCAGCGCAAGTGAACCAAACTATCCAAATTTGCCCTCTCCGCG 450
DB 397 CTCAAGCTCACTTCAAACTTCCGTTAATGCTAATGGAATTAACCTCACTCGTGGGC 456
QY 451 CATACCGCGGTGCGCAAAACCGCTTGGGTGCGCTTAAAGAAATTTG--GGTACTCC 507
DB 457 CACACCGCGGTGCGCAAAACCGCTTGGGTGCGCTAAGCCATGCGCAATTAAGAC 516
QY 508 TCGAATCTAAAGATCTGCACTTGAATCGGTATGATTCAGTGAAGGAGGAAAGG 567
DB 517 CCATCATCACTTTTTCAGCTTAAATAGGAATGATTCAGTGAAGGAGGAACTAACAATAC 576
QY 568 AAACAAACCCCTCTCCGCTGTGGCTTACCTTCCAACTCATTTGACTAGACAAAACG 627
DB 577 ATTAGAACGATCCGATATCTTAACGTTAAACCGGAATCTTTGAGGTGAGCAATACCG 636
QY 628 CCTAATCTGTGATGCTGCTGGGCTTGGTGAACCGCTCGGACCATTTATCCACCG 687
DB 637 GTTGAAGTGTG--GGAACCGGACTCGG--ACGAAAGTGAACACGTGATGCAACA 690
QY 688 TGTGACCTCCCGAGTGAATCAACGAGAGTCTTTTGGGAATGTCAAGGTCCACATGAG 747
DB 691 TCGCACCAAGGAGCTTAACATGAGAGTCTTTTCAAAAGATGTAAAGGCAAGAAAGCC 750
QY 748 CATTTCTGTGCAAGGATTAAGGCAATTTGACATGCTTGAATGATATCAAAAGGAT 807
DB 751 CATTTCTGTGCTGCGGATTAAGGCAATTAAGTGAATGATTTGCGCGGTTT 810
QY 808 AGAGGAGAGATTTCTTATTTGTTGTGAAGATGCTGA--AGAGAGGAGCAATGAGG 864
DB 811 GTTGGATTATGCGCGGTGTGATGTGTAAGATGAGGCAAAAGAAAGTGTGAGATGAGG 870

QY 865 AGATTGCTGAGGACTTGTGTATCATTTTGAAGGCTTATTTGAAGAGAGA 917
DB 871 AGCTTTGAGTGAATTTGTGTTGCTTTCTCAAGTATAGTTTGTGGGGTGA 923

RESULT 13
US-10-381-123-7
Sequence 7, Application US/10381123
Publication No. US20050081263A1
GENERAL INFORMATION:
APPLICANT: E.I. du Pont de Nemours and Company
TITLE OF INVENTION: Chlorophyllases
FILE REFERENCE: B81477 PCT
CURRENT APPLICATION NUMBER: US/10/381,123
PRIOR FILING DATE: 2003-03-21
PRIOR APPLICATION NUMBER: PCT/US01/31059
PRIOR FILING DATE: 2001-10-04
PRIOR APPLICATION NUMBER: 60/238,161
PRIOR FILING DATE: 2000-10-05
NUMBER OF SEQ ID NOS: 30
SOFTWARE: Microsoft Office 97
SEQ ID NO: 7
LENGTH: 1174
TYPE: DNA
ORGANISM: Glycine max
US-10-381-123-7

Query Match 13.3%; Score 151; DB 9; Length 1174;
Best Local Similarity 55.0%; Pred. No. 4.1e-30;
Matches 433; Conservative 0; Mismatches 330; Indels 24; Gaps 6;

QY 143 CTCTAGAGCTTCAACGCTCTCCGCAAGACGCTGTGGTACGCGCGTGAGGAGG 202
DB 97 CATCAATGCTTCTCTCTCACTTCAAAACATTTGTTATCTTACCAACGCTGCTG 156
QY 203 GAGATTAATCGGTGATGATGCTCTCCATGTTACCTTCTCAACCTCTTATTTCTC 262
DB 157 GCTATACCTGTAATTTGTTGTCGATGATTTTCCCTTCCGCAATGATTAATCTG 216
QY 263 AGCTTAATGATGATGCTCTCTCTCAATGCTTCAATCCATGCTCTCTCAATTAATGTA 322
DB 217 AGCTTCAAGGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 276
QY 323 TCGCGGAGCC--AGACCAATGATGATGATGATGATGATGATGATGATGATGATGATG 376
DB 277 GTTACGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 336
QY 377 GGTATC--AGTAGAATTAATCACTTCTTCCAGGCAAGTAAACCAACCTATCCA 433
DB 337 GGTAGCGAGAGAGGCTTCAACCTCTGCTTCCAGAGAAATGTTGAAGCCAAATTTGATA 396
QY 434 AATTGCGCTCTCCGCTCAATGCGCGGTGCAAAACCGGCTTGGCGCTTAAAGA 493
DB 397 AATTGCTTATCAAGTCAAGGCAAGGATGATGATGATGATGATGATGATGATGATGATG 456
QY 494 AATTGGGATCTCTCGAATCTAAAGATCTGACATGATGCGATGATGATGATGATGATG 553
DB 457 ATGCT-----AAACTTAACCTCAAGTTTTCAGCATGATGATGATGATGATGATG 510
QY 554 GACAGGAGAAAGGAGAAACAAACCCCTCTCCGCTGTGGCTTACCTTCAAACTCATTTG 613
DB 511 GCCCATGTAATCTTGGGAAACATTTCTCTTATTTCACTGAGATGCTCCAAATCTTTCA 570
QY 614 ACTTAGCAAAAGCCCTATATCTGTATGATGCTGCGGCTTGTGAAACCGCTCGGAAC 673
DB 571 AATTGAACATA--CCCATTTGTGTAATTTGCACTGGGCTTAAGGCCAAGAGGCTAAT 627
QY 674 CATTAATCCACGCTGAGCACTCCCGAGTGAATCAACGAGATTTCTTTCGGAGATGTC 733
DB 628 TTTTATTTCAACATGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATG 687
QY 734 AAGGTCAAGATGCAATTTGCTGTCGAAGATTAATGAGCAATTTGACATGCTTATGATG 793

Db 688 AACCCCTTGATGACATTTGTTGCACTGATGATGTCACATGACATGATGATGATG 747
Qy 794 ATACAAAAGGATTAAGCG--AAAGTTCTTATTTGTTGTAAGATGTAAG-- 847
Db 748 TGACACCTGGCTTAATGGGTCAATATGTCATAATGTAATGCAAGATGGAAGGTC 807
Qy 848 AGAGAGACCAATGAGGATTCGTTGTTGCACTGTTGTAATGATTTTGAAGCTTAT 907
Db 808 CTAGGACTTATGAAAGAACCGTGGAGGATGTTGTCCTTTTAAGGCAAGT 867
Qy 908 TGAAGG 914
Db 868 TGAATGG 874

RESULT 14
US-10-381-123-17
; Sequence 17, Application US/10381123
; Publication No. US20050081263A1
; GENERAL INFORMATION:
; APPLICANT: E.I. du Pont de Nemours and Company
; TITLE OF INVENTION: Chlorophyllases
; FILE REFERENCE: B01477 PCT
; CURRENT APPLICATION NUMBER: US/10/381,123
; PRIOR FILING DATE: 2003-03-21
; PRIOR APPLICATION NUMBER: PCT/US01/31059
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: 60/238,161
; PRIOR FILING DATE: 2000-10-05
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 17
; LENGTH: 1242
; TYPE: DNA
; ORGANISM: Trifolium aestivum
US-10-381-123-17

Query Match 13.1%; Score 148.2; DB 9; Length 1242;
Best Local Similarity 55.2%; Pred. No. 2.5e-29;
Matches 357; Conservative 0; Mismatches 278; Indels 12; Gaps 3;
Qy 151 GCTTACCGTCTCGCCAAAGCAGCTGTGGTGTGCTAGCCGGTGAAGAGATAT 210
Db 173 GCGGCGCCGACACACCGATCCGGTGTGATCTGCGACCCAAAGATGCAAGACTAC 232
Qy 211 CCGGTGTGATGCTCTTCATGCTTACCTTCTTACAACTCTTATTCAGCTTATG 270
Db 233 CCGGTGTGATGCTCTTCATGCTTACCTTCTTACAACTCTTATTCAGCTTATG 270
Qy 271 TTGATGCTCTTCTTCTGATGCTTATCTTATCTTATCTTATCTTATCTTATCT 327
Db 293 CGGACGCTGATCCACGCGCTTATCTTATCTTATCTTATCTTATCTTATCTTAT 352
Qy 328 GGACGACACATGATGATGATTAATCAACGCGGAGATTAAGATGTTATGATTA 387
Db 353 CTTTGGGTGACGACAGAGATGCGCGCGGACCAAGTGGAGATGCTCCCGAC 412
Qy 388 GGAATTATCACTTCTTCAGCGCAATTAACCAAACTTATCAAACTTATCCCTCTCC 447
Db 413 GGCCTCCCGTCCGCTGCTCCCAAGGCGTGAAGCGGAGCTCTGAAGCTCGCTTGGCC 472
Qy 448 GGCATAGCCCGGTGGCAAAACCGCTTTCGCTTGAAGAAATTTGGTACTCC 507
Db 473 GGCACAGCCGAGAGGCAACGCGCTTCTCCCTGGCTTGGGCAACGCAAGAC-- 529
Qy 508 TCGAATCTAAGATCTGACATTTGATGCTGATTAAGTCAAGTGAAGAGGAAAGG 567
Db 530 ---CAGCTAACCTTCTCGCGCTCATGCGACTCAACCCGCTCGCGGCAAGGAAAGTCC 586
Qy 568 AAACAAACCCCTCTCGGTGGTGGCTTACCTTCAAACTGATTTGACCTAAGCAAAAG 627
Db 587 TCCAGAGCTCAGCCCAAGATCTCACTAAGAGCGGTCTCTTCGCGATG--CGATG 643

Qy 628 CTTATCTTGTGATCGGCTTGGGGCTTGGTGAACCGCTCGAACCCTATTTCCACCG 687
Db 644 CCGGTGTGCTCATTCGCGACCGGGCTCGGGAGAGAGAAATATTTCTTCCCTCC 703
Qy 688 TGTGACCTCCCGAGTGAATCAACCGAGCTTCTTTCGGGAATGTAAGTTCAGCATGG 747
Db 704 TGGCACCCTCAAGAGAGTGAACCAACGCGGATTTCTACCGAGTGGAGCGCGCTGCTAC 763
Qy 748 CATTCGTTGCAAGATTAAGGATTTGATGATGATGATGATGATGATGATGATG 794
Db 764 TACTTGTGACCAAGAGACTACGGGCACTGATGATGATGATGATGATGATGATG 810

RESULT 15
US-10-381-123-9
; Sequence 9, Application US/10381123
; Publication No. US20050081263A1
; GENERAL INFORMATION:
; APPLICANT: E.I. du Pont de Nemours and Company
; TITLE OF INVENTION: Chlorophyllases
; FILE REFERENCE: B01477 PCT
; CURRENT APPLICATION NUMBER: US/10/381,123
; PRIOR FILING DATE: 2003-03-21
; PRIOR APPLICATION NUMBER: PCT/US01/31059
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: 60/238,161
; PRIOR FILING DATE: 2000-10-05
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 9
; LENGTH: 1104
; TYPE: DNA
; ORGANISM: Glycine max
US-10-381-123-9

Query Match 12.8%; Score 145.4; DB 9; Length 1104;
Best Local Similarity 53.8%; Pred. No. 1.4e-28;
Matches 421; Conservative 0; Mismatches 341; Indels 21; Gaps 5;
Qy 143 CTTTGAAGCTTCAACGCTCTCCGCAAGAGCGTGTGATGCTAGCCCGGTGAGAGAG 202
Db 108 CATTCAGTGTCTTCTTCTGACCTCAAAACATTTGTAATTTTACCAACGATTCCTG 167
Qy 203 GAGATTAATCCGGTGTGATGCTCTTCATGCTTATCTTATCTTATCTTATCTTAT 262
Db 168 GCGATACCTCTGAATATGTTGCTGATGCTTTCATCTTCAATTTCTACTACCA 227
Qy 263 AGCTAATGTCATGCTCTTCTGATGCTTCAATCTGATGCTGCTGATGATGATGAT 322
Db 228 AGCTCCTAGCCCAATGATGCTCAATGATGATGATGATGATGATGATGATGATG 287
Qy 323 TCCCGGAGC--AGACAAATGATGATGATTAATCAACGCGGAGATTAAGATGAT 379
Db 288 ATGGGCTTCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 347
Qy 380 TATCAGTGAATTAATCACTTCTTTCAGCGCAAGTGAACCAAACTTATCAAAATTTG 439
Db 348 TACTGAGGAGCTTCAACATTTCTTCAAGAGAACGTTGAAGCAATTTGGAACAACTGG 407
Qy 440 CCGTCTCCCGGCAATGCGCGGAGGCAAAACGCGTTTGGCGGCTTAAAGAAATTTG 489
Db 408 TTTATGAGTCAACAGTGGGAGGAGAAATCTTAATTTGCTGGCTCT--TGCTC 461
Qy 500 GGTACTCTCGAATCTAAGATCTGACATTTGATGCTGATTAAGTCAAGTGAAGACAG 559
Db 462 ATGCAAAATTAATCTCAAGTTTCAAGCACTTGTAGCAATAGACCTGTGGCTGGCAAT 521
Qy 560 GGAAGGAGAAACAAACCCCTCTCGCGTGTGGCTTACCTTCAAACTGATTTGACCTAG 619
Db 522 CTAAATATTTAGAAACGCTCTCATATTTCTCACTGCAAGCAACGCTCTTGAATTT-- 579
Qy 620 ACAAAACGCTATATTTGATGCTTGGGAGCTTGGTGAACCGCTCGGAACCAATTA 679

Db 580 -GAAATGCCAGTTGAAATTTGGCACTGATGGGCCAGAGAAGCTTAATTGTGTA 638
Qy 680 TCCCAACCGTGTGCACCTCCCCGAGTGAATCACCGAGATTCTTCCGGAATGTCAAGTC 739
Db 639 CTCACCCGTGTCTCTGTATGGGGTGAACTATPAGGAGTTCTTCAACGAGTGCMAAACC 698
Qy 740 CAGCATGGCATTTCGTTGGCAAGATTATGGGCATTGGACATGGCTTGATGATACAA 799
Db 699 CTTGTGCTAAATTTGTGTAGCAAAGTATGTCACTGACATGACATGTTGAATGATGACAC 758
Qy 800 AAGGATTAGAGGA--AGAGTTCTTATTGTTGTGAAGATG-----TGAAGAGA 850
Db 759 CAGGGCTAATGGGACATGTGTCAAGTGTATGTATGATPAGAAATGGACGAGGTCCTA 818
Qy 851 GGAGACCAATGAGAGATTGTTGTGAGACTTGTGTATCATTTTGAAGGCTTATTGG 910
Db 819 GGGACTTGATGAAAGAACCACTGGAGGGTGTGTGTCTTCTTGAGGGCACAAATTGA 878
Qy 911 AAG 913
Db 879 ATG 881

Search completed: March 20, 2006, 14:42:22
Job time : 1056.94 secs

Db 355 GTATGTTTATCATATCATAGGCTATTAAGATCATATCTCT 394

RESULT 2

US-10-330-773-37
; Sequence 37, Application US/10330773
; Publication No. US20060040262A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc Malandro
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer
; FILE REFERENCE: 529452001300
; CURRENT APPLICATION NUMBER: US/10/330,773
; CURRENT FILING DATE: 2002-12-27
; NUMBER OF SEQ ID NOS: 981
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 37
; LENGTH: 135462
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(135462)
; OTHER INFORMATION: n = A,T,C or G
US-10-330-773-37

Query Match 3.8%; Score 43; DB 7; Length 135462;
Best Local Similarity 52.5%; Pred. No. 31;
Matches 117; Conservative 0; Mismatches 105; Indels 1; Gaps 1;

QY 9 TAAAGAAAAGAAAAGAACTAATAAGAAACAAAAAAATGCTCTTCATCAAGAAAGC 68
Db 124353 TAAAAAAGAAAGAAACAAAACAAAAAACCAGACCTGCTTCCCAAGTCTTA 124412
QY 69 CTTTGAAGATGCGCAATATACAAATCTTAACTTGAAGCTATCATCTGTTGCTG 128
Db 124413 CTTTAAAGCTTTAGATGAGAACCAATGCTGCTCATAGTAGAATGTGGTGT -GGAGCTC 124471
QY 129 CAAATTAACACCGTCTTCTAGAGCTTCAACCGTCTCCGCAAGCAGCTGTGGTGTAC 188
Db 124472 CAAAAGGCCCGACGTGGCGAGTAGTACTGTGACCCGGCCTTACAGCTGTGTCTGCTTT 124531
QY 189 GCCGGTGAAGAGAGATTATCCGGTGTGATGCTCTCCAT 231
Db 124532 CCTGTGTCGTCACGAGTTTGTTCCTCGAGTCATGTTTACCGT 124574

RESULT 3

US-09-925-065A-343968
; Sequence 343968, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 343968
; LENGTH: 561
; TYPE: DNA

; ORGANISM: Homo sapiens
US-09-925-065A-343968

Query Match 3.7%; Score 41.8; DB 6; Length 561;
Best Local Similarity 53.5%; Pred. No. 10;
Matches 85; Conservative 1; Mismatches 73; Indels 0; Gaps 0;

QY 977 AAGAGTTGAGGTATCATCATGTAACATTAAGTTTCTTTAGGGGCTGTTTCTATTGT 1036
Db 101 AATGATGATATATATCAAAATTAAGAGATTCATTAAGATTCATTAATTTATTTT 160
QY 1037 CAATATCATCAGCTTTTGTGCTTATGTTTACAACTTATATGTACATCTTTAAG 1096
Db 161 ACTATTCCTCTCTTTGTATGAAAGATCTTAATTAATATATTAATAAAGCATTTCT 220
QY 1097 TCACCTCTTGTCTTACAAAAAATTAATAAATAA 1135
Db 221 CCAATTCCTGTTAATTAATAAATAAAGAGAGAA 259

RESULT 4

US-09-925-065A-343967
; Sequence 343967, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 343967
; LENGTH: 561
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-343967

Query Match 3.6%; Score 40.8; DB 6; Length 561;
Best Local Similarity 52.4%; Pred. No. 17;
Matches 87; Conservative 1; Mismatches 78; Indels 0; Gaps 0;

QY 970 GAAATTCAGAGTTTGAAGTTATCATGTAAACATTAAGTTTCTTTAGGGGCTGTTT 1029
Db 94 RAAAGAAAATATGATGATATATCAATTAAGATTCATTAAGATTCATTAATTTATTT 153
QY 1030 CTATTCATATATATCATGAGCTTTTGTGCTTATGTTTACAACTTATATGTACACT 1089
Db 154 TTATTTTACTAATTCCTCTCTTTGTATGAAAGATCTTAATTAATAATATTAATAAAG 213
QY 1090 CTTTAAGTCACCTTTGCTTACAAAAAATTAATAAATAA 1135
Db 214 CATTCCTCAATTCCTGTTAATTAATAAATAAAGAGAGAA 259

RESULT 5

US-10-947-249-91
; Sequence 91, Application US/10947249
; Publication No. US20050287541A1
; GENERAL INFORMATION:
; APPLICANT: Akira NAKAGAWARA
; APPLICANT: Miki OHIRA

APPLICANT: Shin ISHII
APPLICANT: Takeshi GOTO
APPLICANT: Hiroyuki KUBO
APPLICANT: Takahiro HIRATA
APPLICANT: Yasuko YOSHIDA
APPLICANT: Saichi YAMADA
TITLE OF INVENTION: Microarray for Predicting the Prognosis of Neuroblastoma and Meth
TITLE OF INVENTION: Predicting the Prognosis of Neuroblastoma
FILE REFERENCE: 117007
CURRENT APPLICATION NUMBER: US/10/947,249
CURRENT FILING DATE: 2004-09-23
PRIOR APPLICATION NUMBER: US 60/505,614
PRIOR APPLICATION NUMBER: 2003-09-25
NUMBER OF SEQ ID NOS: 200
SOFTWARE: PatentIn version 3.1
SEQ ID NO 91
LENGTH: 1721
TYPE: DNA
ORGANISM: Homo sapiens
US-10-947-249-91

Query Match 3.6%; Score 40.6; DB 8; Length 1721;
Best Local Similarity 58.8%; Pred. No. 27;
Matches 70; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY 1017 GGGGCTGTTTCTTATGTCATATATCATGCTTTGCTTATGCTTTTACAACTT 1076
DB 1584 GGGGAAGATTGGGCTTGTCCCAACACAGCTTGTGCTGCTACTGTAATCTGTACA 1643
QY 1077 ATATTGACAACTGTTTAAGTCACCTCTTGGCTTACAAAAA 1135
DB 1644 ACTGTTTCTGACCACTTAATAGCTGTTGTAAGTAAAAA 1702

RESULT 6
US-11-122-329-76
Sequence 76, Application US/11122329
Publication No. US20060019272A1
GENERAL INFORMATION:
APPLICANT: Geraci, Mark
APPLICANT: Bull, Todd
APPLICANT: Voelkel, Norbert
APPLICANT: Coldren, Chris
TITLE OF INVENTION: Diagnosis of Disease and Monitoring of Therapy Using Gene
TITLE OF INVENTION: Expression Analysis of Peripheral Blood Cells
FILE REFERENCE: 2848-54
CURRENT APPLICATION NUMBER: US/11/122,329
CURRENT FILING DATE: 2005-05-03
PRIOR APPLICATION NUMBER: 60/568,129
PRIOR FILING DATE: 2004-05-03
NUMBER OF SEQ ID NOS: 128
SOFTWARE: PatentIn version 3.3
SEQ ID NO 76
LENGTH: 1721
TYPE: DNA
ORGANISM: Homo sapiens
US-11-122-329-76

Query Match 3.6%; Score 40.6; DB 12; Length 1721;
Best Local Similarity 58.8%; Pred. No. 27;
Matches 70; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY 1017 GGGGCTGTTTCTTATGTCATATATCATGCTTTGCTTATGCTTTTACAACTT 1076
DB 1584 GGGGAAGATTGGGCTTGTCCCAACACAGCTTGTGCTGCTGTAATCTGTACA 1643
QY 1077 ATATTGACAACTGTTTAAGTCACCTCTTGGCTTACAAAAA 1135
DB 1644 ACTGTTTCTGACCACTTAATAGCTGTTGTAAGTAAAAA 1702

RESULT 7
US-11-096-568A-12837

Sequence 12837, Application US/11096568A
Publication No. US20060048240A1
GENERAL INFORMATION:
APPLICANT: Alexandrov, Nikolai et al.
TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
FILE REFERENCE: 2750-1592PUS2
CURRENT APPLICATION NUMBER: US/11/096,568A
CURRENT FILING DATE: 2005-04-01
NUMBER OF SEQ ID NOS: 34471
SEQ ID NO 12837
LENGTH: 1842
TYPE: DNA
ORGANISM: Trilicium aestivum
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)-(1842)
OTHER INFORMATION: Ceres Seq. ID no. 14303706
US-11-096-568A-12837

Query Match 3.6%; Score 40.4; DB 9; Length 1842;
Best Local Similarity 65.6%; Pred. No. 30;
Matches 59; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

QY 1046 CAGCTTTGCTTATGCTTTTACAACTTATATGTAACCTTTAAGTCACCTT 1105
DB 1702 CAGTATATCTTGTATGATGATGATATATATATATGTAATGTAACCTTATATGTCACCTT 1761
QY 1106 TGCTTCAAAAAA 1135
DB 1762 TGTGTTGAAAAA 1791

RESULT 8
US-09-925-065A-84331
Sequence 84331, Application US/09925065A
Publication No. US20040181048A1
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single
FILE REFERENCE: 108827.135
CURRENT APPLICATION NUMBER: US/09/925,065A
CURRENT FILING DATE: 2001-08-08
PRIOR APPLICATION NUMBER: US 60/243,096
PRIOR FILING DATE: 2000-10-24
PRIOR APPLICATION NUMBER: US 60/252,147
PRIOR FILING DATE: 2000-11-20
PRIOR APPLICATION NUMBER: US 60/250,092
PRIOR FILING DATE: 2000-11-30
PRIOR APPLICATION NUMBER: US 60/261,766
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: US 60/289,846
PRIOR FILING DATE: 2001-05-09
NUMBER OF SEQ ID NOS: 957086
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 84331
LENGTH: 652
TYPE: DNA
ORGANISM: Homo sapiens
US-09-925-065A-84331

Query Match 3.5%; Score 40.2; DB 6; Length 652;
Best Local Similarity 84.9%; Pred. No. 24;
Matches 45; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1083 TAAACTCTTATGTCACCTCTTGTCTTACAAAAA 1135
DB 132 TATATTGTTAAAGTCACCTTCACTTAAAAA 184

RESULT 9
US-09-925-065A-84332

```
; Sequence 84332, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 84332
; LENGTH: 652
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-84332
```

```
Query Match 3.5%; Score 40.2; DB 6; Length 652;
Best Local Similarity 84.9%; Pred. No. 24;
Matches 45; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
```

```
Qy 1083 TACACCTTTTAACTGCTCTTGTTCACAAAAA 1135
Db 132 TATATTGTTAAAGTCACTCTTCACTTAAAAA 184
```

```
RESULT 10
US-11-050-857-1056
; Sequence 1056, Application US/11050857
; Publication No. US20060040278A1
; GENERAL INFORMATION:
; APPLICANT: Compugen Ltd
; TITLE OF INVENTION: NOVEL NUCLEOTIDE AND AMINO ACID SEQUENCES, AND ASSAYS AND METHODS
; FILE REFERENCE: 1847.1005
; CURRENT APPLICATION NUMBER: US/11/050,857
; CURRENT FILING DATE: 2005-01-27
; NUMBER OF SEQ ID NOS: 1150
; SEQ ID NO 1056
; LENGTH: 2201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-050-857-1056
```

```
Query Match 3.5%; Score 40.2; DB 9; Length 2201;
Best Local Similarity 70.1%; Pred. No. 35;
Matches 54; Conservative 0; Mismatches 23; Indels 0; Gaps 0;
```

```
Qy 1059 TTAAGTTTACAACTTATATTGTACAACTTTAAGTCACTCTTGTCTTACAAAAA 1118
Db 2122 TTAAGTTTCAACAATTTCATTCGCTTCCCTTCAATAAAGTTGTCATTCACAAA 2181
Qy 1119 AAAAAAAAAAAAAAAAAA 1135
Db 2182 AAAAAAAAAAAAAAAAAA 2198
```

```
RESULT 11
US-11-051-720-1711
; Sequence 1711, Application US/11051720
; Publication No. US20060046257A1
; GENERAL INFORMATION:
; APPLICANT: Compugen Ltd
```

```
; TITLE OF INVENTION: NOVEL NUCLEOTIDE AND AMINO ACID SEQUENCES, AND ASSAYS AND METHODS
; FILE REFERENCE: 1847.1002
; CURRENT APPLICATION NUMBER: US/11/051,720
; CURRENT FILING DATE: 2005-01-27
; NUMBER OF SEQ ID NOS: 1780
; SEQ ID NO 1711
; LENGTH: 2201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-051-720-1711
```

```
Query Match 3.5%; Score 40.2; DB 9; Length 2201;
Best Local Similarity 70.1%; Pred. No. 35;
Matches 54; Conservative 0; Mismatches 23; Indels 0; Gaps 0;
```

```
Qy 1059 TTAAGTTTACAACTTATATTGTACAACTTTAAGTCACTCTTGTCTTACAAAAA 1118
Db 2122 TTAAGTTTCAACAATTTCATTCGCTTCCCTTCAATAAAGTTGTCATTCACAAA 2181
```

```
Qy 1119 AAAAAAAAAAAAAAAAAA 1135
Db 2182 AAAAAAAAAAAAAAAAAA 2198
```

```
RESULT 12
US-11-043-806-516
; Sequence 516, Application US/11043806
; Publication No. US20060051774A1
; GENERAL INFORMATION:
; APPLICANT: Compugen Ltd
; TITLE OF INVENTION: Novel Nucleotide and Amino Acid Sequences, and Assays and Methods
; FILE REFERENCE: 1847.1003
; CURRENT APPLICATION NUMBER: US/11/043,806
; CURRENT FILING DATE: 2005-01-27
; NUMBER OF SEQ ID NOS: 575
; SEQ ID NO 516
; LENGTH: 2201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-043-806-516
```

```
Query Match 3.5%; Score 40.2; DB 9; Length 2201;
Best Local Similarity 70.1%; Pred. No. 35;
Matches 54; Conservative 0; Mismatches 23; Indels 0; Gaps 0;
```

```
Qy 1059 TTAAGTTTACAACTTATATTGTACAACTTTAAGTCACTCTTGTCTTACAAAAA 1118
Db 2122 TTAAGTTTCAACAATTTCATTCGCTTCCCTTCAATAAAGTTGTCATTCACAAA 2181
```

```
Qy 1119 AAAAAAAAAAAAAAAAAA 1135
Db 2182 AAAAAAAAAAAAAAAAAA 2198
```

```
RESULT 13
US-11-043-590-46
; Sequence 46, Application US/11043590
; Publication No. US20050277156A1
; GENERAL INFORMATION:
; APPLICANT: Compugen Ltd
; TITLE OF INVENTION: Novel Brain Natriuretic Peptide Variants and Methods of use thereof
; FILE REFERENCE: 1847.1011
; CURRENT APPLICATION NUMBER: US/11/043,590
; CURRENT FILING DATE: 2005-01-27
; NUMBER OF SEQ ID NOS: 47
; SEQ ID NO 46
; LENGTH: 2201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-043-590-46
```

```

RESULT 15
US-10-750-185-31978
: Sequence 31978, Application US/10750185
: Publication No. US20050260603A1
: GENERAL INFORMATION:
: APPLICANT: MMT GENOMICS, INC.
: APPLICANT: DENISE, Sue K.
: APPLICANT: KERR, Richard
: APPLICANT: ROSENFELD, David
: APPLICANT: HOLM, Tom
: APPLICANT: BATES, Stephen
: APPLICANT: PANTIN, Dennis
: TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
: FILE REFERENCE: MW1100-2

```

Accession	Sequence	Position
Oy	GATTAGAGGGGAAGGTTCTTAATGTGTTGTGTAAGATGGTGAAGAGGAAACCAATGAG	863
Dp	542 GACAAATGGGGCTATTTTGGCTGTCTTTCTTTGGTGCAATGCAATGCTCAAGATATCA	601
Oy	864 GAGATTCGTGGTGGACTGTGTATCATTTTGAAGGCTTATTTGAAGAGATGATCG	923
Dp	602 AGTTCGTGACACGTCTCTCGTTCTGTGACACTGTAATAAATAATTTTTTCCCGAGCC	661
Oy	924 TGAATTGATTAAGATCAAAAGATGGGTGTCAAGAGATGTTCCGTTGAATTCAAAGATT	983
Dp	662 TTAAATTCCTTCATCTGTAACATGGGTAAATATGAGAGATTAATTAATATGAACTCCGAATTA	721
Oy	984 TGAGCTATCATGTAAACATTAAGTTTTCTTTAAGGGCTGTTTTCTATTGTCATATTC	1043
Dp	722 TGAGGATTTAAGTCMAAGATGACCCCTGGGGGCTTAAATCTGTGTTCCGTGTGGCTC	781
Oy	1044 ATCACTTTTGTGGCTATGAGTTTACAACCTTATTTATGTACAACCTCTTAAAGTACACTC	1103
Dp	782 ACCAGATTCAGTGTCCTTATGAGCTTGCCCAATTCAGAAAGTCTTTCTCTCAATGAAAAACC	841
Oy	1104 TTTCG 1107	
Dp	842 TTAG 845	

Search completed: March 20, 2006, 14:23:27
Job time : 454.902 secs

THIS PAGE LEFT BLANK

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 20, 2006, 13:57:59 ; Search time 226.218 Seconds
(without alignments)
8918.522 Million cell updates/sec

Title: US-10-634-548-19

Sequence: 1 aaaaaaagctaaagaaagaa.....aaaaaaaaaaaaaaaa 1135

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA:*

- 1: /cgn2_6/ptodate/1/ina/1_COMB.seq:*
- 2: /cgn2_6/ptodate/1/ina/5_COMB.seq:*
- 3: /cgn2_6/ptodate/1/ina/6A_COMB.seq:*
- 4: /cgn2_6/ptodate/1/ina/6B_COMB.seq:*
- 5: /cgn2_6/ptodate/1/ina/H_COMB.seq:*
- 6: /cgn2_6/ptodate/1/ina/PTUS_COMB.seq:*
- 7: /cgn2_6/ptodate/1/ina/PP_COMB.seq:*
- 8: /cgn2_6/ptodate/1/ina/RE_COMB.seq:*
- 9: /cgn2_6/ptodate/1/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	44.2	3.9	1979	US-09-461-325-30	Sequence 30, Appl
2	44.2	3.9	1979	US-10-012-542-30	Sequence 30, Appl
3	44.2	3.9	1979	US-10-115-123-30	Sequence 30, Appl
4	43	3.8	198	US-08-330-108-16	Sequence 16, Appl
5	43	3.8	198	PCT-US92-10087-16	Sequence 16, Appl
6	42.8	3.8	1637	US-08-270-767-11283	Sequence 11283, A
7	42.4	3.7	991	US-08-924-747-25	Sequence 25, Appl
8	42.4	3.7	991	US-09-247-3738-25	Sequence 25, Appl
9	42.4	3.7	991	US-09-236-715-25	Sequence 25, Appl
10	42	3.7	3265	US-09-832-129-13	Sequence 13, Appl
11	41.6	3.7	640681	US-09-790-988-1	Sequence 1, Appl
12	40.4	3.6	2186	US-10-033-301-23	Sequence 23, Appl
13	40	3.5	1278	US-08-008-687A-11	Sequence 11, Appl
14	40	3.5	2862	US-08-148-209A-1	Sequence 1, Appl
15	39.6	3.5	1380	US-09-881-556A-1	Sequence 1, Appl
16	39.6	3.5	7260	US-09-917-254-31	Sequence 31, Appl
17	39.6	3.5	7260	US-09-919-497-24	Sequence 24, Appl
18	39.4	3.5	1722	US-09-134-001C-1976	Sequence 1976, Ap
19	39.4	3.5	2007	US-08-747-221B-36	Sequence 36, Appl
20	39.4	3.5	2007	US-08-747-221B-38	Sequence 36, Appl
21	39.4	3.5	2007	US-09-005-051-36	Sequence 36, Appl
22	39.4	3.5	2007	US-09-005-051-38	Sequence 38, Appl
23	39.4	3.5	2007	US-09-403-942F-36	Sequence 36, Appl
24	39.4	3.5	2007	US-09-403-942F-38	Sequence 38, Appl

25	39.2	3.5	1141	US-09-806-708B-22	Sequence 22, Appl
26	39.2	3.5	1582	US-08-545-196B-10	Sequence 10, Appl
27	39.2	3.5	1582	US-08-545-196B-12	Sequence 12, Appl
28	39.2	3.5	1737	US-08-202-056-4	Sequence 4, Appl
29	39.2	3.5	1737	US-08-076-093A-3	Sequence 3, Appl
30	39.2	3.5	1737	US-08-101-265-3	Sequence 3, Appl
31	39.2	3.5	1737	US-08-284-586-3	Sequence 3, Appl
32	39.2	3.5	1737	US-08-805-478-3	Sequence 3, Appl
33	39.2	3.5	1737	US-08-802-627A-3	Sequence 3, Appl
34	39.2	3.5	1737	US-08-801-238-3	Sequence 3, Appl
35	39.2	3.5	1737	US-08-801-228-3	Sequence 3, Appl
36	39.2	3.5	1737	US-09-104-296-3	Sequence 3, Appl
37	39.2	3.5	1737	PCT-US94-06380-2	Sequence 2, Appl
38	39.2	3.5	4709	US-08-956-171E-587	Sequence 587, App
39	39.2	3.5	4709	US-08-781-986A-587	Sequence 587, App
40	39.2	3.5	77586	US-09-949-016-13220	Sequence 13220, A
41	39.2	3.5	77586	US-09-949-016-13221	Sequence 13221, A
42	39	3.4	1664976	US-08-916-421B-1	Sequence 1, Appl
43	39	3.4	1664976	US-09-692-570-1	Sequence 1, Appl
44	38.8	3.4	213456	US-09-820-007-3	Sequence 3, Appl
45	38.6	3.4	1939	US-09-719-083A-3	Sequence 3, Appl

ALIGNMENTS

```
RESULT 1
US-09-461-325-30
; Sequence 30, Application US/09461325A
; Patent No. 6475753
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 94 Human Secreted Proteins
; FILE REFERENCE: P2029P1
; CURRENT APPLICATION NUMBER: US/09/461,325A
; EARLIER FILING DATE: 1999-12-14
; EARLIER APPLICATION NUMBER: PCT/US99/13418
; EARLIER FILING DATE: 1999-06-15
; EARLIER APPLICATION NUMBER: 60/089,507
; EARLIER FILING DATE: 1998-06-16
; EARLIER APPLICATION NUMBER: 60/089,508
; EARLIER FILING DATE: 1998-06-16
; EARLIER APPLICATION NUMBER: 60/089,509
; EARLIER FILING DATE: 1998-06-16
; EARLIER APPLICATION NUMBER: 60/089,510
; EARLIER FILING DATE: 1998-06-16
; EARLIER APPLICATION NUMBER: 60/090,112
; EARLIER FILING DATE: 1998-06-22
; EARLIER APPLICATION NUMBER: 60/090,113
; EARLIER FILING DATE: 1998-06-22
; NUMBER OF SEQ ID NOS: 532
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 30
; LENGTH: 1979
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (968)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-461-325-30

Query Match 3.9%; Score 44.2; DB 3; Length 1979;
Best Local Similarity 53.8%; Pred. No. 0.084;
Matches 91; Conservative 0; Mismatches 78; Indels 0; Gaps 0;

QY 967 GTTGAATTCACAGCTTGAAGCTTATCATGTAACATAGTTTCTTAAAGGCGTGGT 1026
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1802 GTTCAAGCTTACTACTATGTAATCAATCATATATATATATATATATATATATAT 1861
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1027 TTTCATATGTCATATATCATGCTTTGTGCTTATGTTTACAACTTATTTGTACA 1086
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1862 ATTTATGTTTATGATCATATATATTTGGGCATATATCTTGTGGATTAAATTAATA 1921
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

QY 1087 ACTCTTAAGTCACCTCTTGTCTACAAAAA 1135
DB 1922 CACTTATATTTTCATGAACCTAAAAA 1970

RESULT 2

US-10-012-542-30

Sequence 30, Application US/10012542

Patent No. 667741

GENERAL INFORMATION:

APPLICANT: Ruben et al.

TITLE OF INVENTION: 94 Human Secreted Proteins

FILE REFERENCE: P2029PI

CURRENT APPLICATION NUMBER: US/10/012,542

CURRENT FILING DATE: 2001-12-12

PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/461,325

PRIOR FILING DATE: EARLIER FILING DATE: 1999-12-14

PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,507

PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16

PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,508

PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16

PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,509

PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16

PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,510

PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16

PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/090,112

PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-22

PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/090,113

PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-22

NUMBER OF SEQ ID NOS: 532

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 30

LENGTH: 1979

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: SITE

LOCATION: (968)

OTHER INFORMATION: n equals a,t,g, or c

US-10-012-542-30

Query Match 3.9%; Score 44.2; DB 3; Length 1979;

Best Local Similarity 53.8%; Pred. No. 0.084;

Matches 91; Conservative 0; Mismatches 78; Indels 0; Gaps 0;

QY 967 GTTGAATTCAAGCTTTGAGGTATCATGTAACATTAAGTTTCTTTAGGGGCTGTT 1026

DB 1802 GTTCAAGCTGCTATTGGTATATCATCTAATTTATATATCTCCAGGCCCTGA 1861

QY 1027 TTTCTATTGTCAATATCATGCTTTGTGCTTATGTTTACAACTTATATTTGACA 1086

DB 1862 ATTTATGTTGATGATCATATATATTTGGCATATATCTTTGGGATTGAAATAATAA 1921

QY 1087 ACTCTTAAGTCACCTCTTGTCTACAAAAA 1135

DB 1922 CACTTATATTTTCATGAACCTAAAAA 1970

RESULT 3

US-10-115-123-30

Sequence 30, Application US/10115123

Patent No. 6774216

GENERAL INFORMATION:

APPLICANT: Ruben et al.

TITLE OF INVENTION: 94 Human Secreted Proteins

FILE REFERENCE: P2029G30AP1D2

CURRENT APPLICATION NUMBER: US/10/115,123

CURRENT FILING DATE: 2002-04-04

PRIOR APPLICATION NUMBER: PCT/US99/13418

PRIOR FILING DATE: 1999-06-15

PRIOR APPLICATION NUMBER: 60/089,507

PRIOR FILING DATE: 1998-06-16

PRIOR APPLICATION NUMBER: 60/089,508
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089,509
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089,510
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/090,112
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090,113
PRIOR FILING DATE: 1998-06-22
NUMBER OF SEQ ID NOS: 532
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 30
LENGTH: 1979
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (968)
OTHER INFORMATION: n equals a,t,g, or c
US-10-115-123-30

Query Match 3.9%; Score 44.2; DB 3; Length 1979;

Best Local Similarity 53.8%; Pred. No. 0.084;

Matches 91; Conservative 0; Mismatches 78; Indels 0; Gaps 0;

QY 967 GTTGAATTCAAGCTTTGAGGTATCATGTAACATTAAGTTTCTTTAGGGGCTGTT 1026

DB 1802 GTTCAAGCTGCTATTGGTATATCATCTAATTTATATATCTCCAGGCCCTGA 1861

QY 1027 TTTCTATTGTCAATATCATGCTTTGTGCTTATGTTTACAACTTATATTTGACA 1086

DB 1862 ATTTATGTTGATGATCATATATTTGGCATATATCTTTGGGATTGAAATAATAA 1921

QY 1087 ACTCTTAAGTCACCTCTTGTCTACAAAAA 1135

DB 1922 CACTTATATTTTCATGAACCTAAAAA 1970

RESULT 4

US-08-330-108-16

Sequence 16, Application US/08330108

Patent No. 5795752

GENERAL INFORMATION:

APPLICANT: Smith, Kendall A.

TITLE OF INVENTION: IL-2-Stimulated Gene

NUMBER OF SEQUENCES: 20

CORRESPONDENCE ADDRESS:

ADDRESS: Lohive & Cockfield

STREET: 60 State Street

CITY: Boston

STATE: Massachusetts

COUNTRY: U.S.A.

ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: ASCII TEXT

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/330,108

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/104,736

FILING DATE:

APPLICATION NUMBER: US/07/796,066

FILING DATE:

INFORMATION FOR SEQ ID NO: 16:

SEQUENCE CHARACTERISTICS:

LENGTH: 198 base pairs

TYPE: nucleic acid

```
STRANDEDNESS: single stranded
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: no
ANTI-SENSE: no
ORIGINAL SOURCE:
ORGANISM: human
CELL TYPE: T-cell blast
IMMEDIATE SOURCE:
LIBRARY:
CLONE: 8D4-T3
US-08-330-108-16

Query Match
Best Local Similarity 60.9%; Score 43; DB 2; Length 198;
Matches 70; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

Qy 1021 CTGGTTTCTATGTGCAATATCATCAGCTTTGTGCTTATGTTTACAACTTATAT 1080
Db 63 CTGTTTATTTTGGTATTAAGTCGTGCTTTATTTGTAAGCTTATTAATATATAT 122

Qy 1081 TGTACAACTCTTTAAGTACCTCTTGTCTTACAAAACAAAAAAAAAAAAA 1135
Db 123 TATATTAATATATATTAAGAAAGAAATGTTTCAGAAAAA 177

RESULT 5
PCT-US92-10087-16
Sequence 16, Application PC/RUS9210087
GENERAL INFORMATION:
APPLICANT: Smith, Kendall A.
TITLE OF INVENTION: IL-2-Stimulated Gene
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lathive & Cockfield
STREET: 60 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII TEXT
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/10087
FILING DATE: 19921118
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/796,066
FILING DATE: 20-NOVEMBER-1991
ATTORNEY/AGENT INFORMATION:
NAME: DeConti, Giulio A. Jr.
REGISTRATION NUMBER: 31,503
REFERENCE/DOCKET NUMBER: DCI-028BC
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-227-7400
TELEFAX: 617-227-5941
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 198 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single stranded
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: no
ANTI-SENSE: no
ORIGINAL SOURCE:
ORGANISM: human
CELL TYPE: T-cell blast
IMMEDIATE SOURCE:
```

```
LIBRARY:
CLONE: 8D4-T3
PCT-US92-10087-16

Query Match
Best Local Similarity 60.9%; Score 43; DB 6; Length 198;
Matches 70; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

Qy 1021 CTGGTTTCTATGTGCAATATCATCAGCTTTGTGCTTATGTTTACAACTTATAT 1080
Db 63 CTGTTTATTTTGGTATTAAGTCGTGCTTTATTTGTAAGCTTATTAATATATAT 122

Qy 1081 TGTACAACTCTTTAAGTACCTCTTGTCTTACAAAACAAAAAAAAAAAAA 1135
Db 123 TATATTAATATATTAAGAAAGAAATGTTTCAGAAAAA 177

RESULT 6
US-09-270-767-11283/c
Sequence 11283, Application US/09270767
Patent No. 6703491
GENERAL INFORMATION:
APPLICANT: Homburger et al.
TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
FILE REFERENCE: File Reference: 7326-094
CURRENT FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 62517
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 11283
LENGTH: 1637
TYPE: DNA
ORGANISM: Drosophila melanogaster
US-09-270-767-11283

Query Match
Best Local Similarity 57.5%; Score 42.8; DB 3; Length 1637;
Matches 77; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

Qy 997 TAAACATTAAGTTTCTTTAGGGGCTGTTTCTATGTCATATCATCAGTTTGT 1056
Db 135 TCAACACTGAGAACATTTCGATTAGCATTTTCCCTTCATATCAATCAATTTGT 76

Qy 1057 GCTTATGTTTTCACAACTTATATGTGACACTTTTAACTCTTTGCTTACAAA 1116
Db 75 AACTATGATTTTGAACAAAAATATATAAATATATATAGTCCTAATGACCAACAAA 16

Qy 1117 AAAAAAAAAAAAAA 1130
Db 15 AAAAAAAAAAAAAA 2

RESULT 7
US-08-924-747-25
Sequence 25, Application US/08924747
Patent No. 6063570
GENERAL INFORMATION:
APPLICANT: MCGONIGLE, BRIAN
APPLICANT: O'KEEFE, DANIEL
TITLE OF INVENTION: SOYBEAN GLUTATHIONE-S-TRANSFERASE
TITLE OF INVENTION: ENZYMES
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: E.I. DU PONT DE NEMOURS AND COMPANY
STREET: 1007 MARKET STREET
CITY: WILMINGTON
STATE: DELAWARE
COUNTRY: UNITED STATES OF AMERICA
ZIP: 19898
COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE, 3.50 INCH
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: MICROSOFT WORD FOR WINDOWS 95
```

SOFTWARE: MICROSOFT WORD VERSION 7.0A
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/924,747
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: FLOYD, LINDA AXAMETHY
REGISTRATION NUMBER: 33,692
REFERENCE/DOCKET NUMBER: CL-1108
TELECOMMUNICATION INFORMATION:
TELEPHONE: 302-892-8112
TELEFAX: 302-773-0164
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 991 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
TISSUE TYPE: SOYBEAN
IMMEDIATE SOURCE:
CLONE: SSM.PK0067.G5
US-08-924-747-25

Query Match 3.7%; Score 42.4; DB 3; Length 991;
Best Local Similarity 54.5%; Pred. No. 0.19;
Matches 85; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

QY 980 AGTTGAGGTTATCATGTAACATAAGTTTCTTTAGGGCGCTGTTTCTATTGTCAA 1039
DB 800 AGTATGATTTTGTGGGAAACAATTATCTTGTGTGAGCAAGATGTTCTGTTTAA 859
1040 TATCATCAGCTTTGTGCTTATGCTTTTACAACTTATATGTACAACCTTTAAGTCA 1099
DB 860 TTTAATGACGCTGATTTGGTATGGCTATTTTAATTTTAACATAAAAAAGTGT 919
QY 1100 CCTCTTGGCTTACAAAAA 1135
DB 920 CAGTTTAAAAA 955

RESULT 8
US-09-247-373B-25
Sequence 25, Application US/09247373B
Patent No. 6168954
GENERAL INFORMATION:
APPLICANT: MCGONIGLE, BRIAN
APPLICANT: O'KEEFE, DANIEL
TITLE OF INVENTION: SOYBEAN GLUTATHIONE-S-TRANSFERASE ENZYMES
FILE REFERENCE: CL-1108-A
CURRENT APPLICATION NUMBER: US/09/247,373B
CURRENT FILING DATE: 1999-02-10
PRIOR APPLICATION NUMBER: 08/924,747
PRIOR FILING DATE: 1997-09-05
NUMBER OF SEQ ID NOS: 56
SOFTWARE: Microsoft Office 97
SEQ ID NO 25
LENGTH: 991
TYPE: DNA
ORGANISM: SOYBEAN
US-09-247-373B-25

Query Match 3.7%; Score 42.4; DB 3; Length 991;
Best Local Similarity 54.5%; Pred. No. 0.19;
Matches 85; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

QY 980 AGTTGAGGTTATCATGTAACATAAGTTTCTTTAGGGCGCTGTTTCTATTGTCAA 1039
DB 800 AGTATGATTTTGTGGGAAACAATTATCTTGTGTGAGCAAGATGTTCTGTTTAA 859

QY 1040 TATCATCAGCTTTGTGCTTATGCTTTTACAACTTATATGTACAACCTTTAAGTCA 1099
DB 860 TTTAATGACGCTGATTTGGTATGGCTATTTTAATTTTAACATAAAAAAGTGT 919
QY 1100 CCTCTTGGCTTACAAAAA 1135
DB 920 CAGTTTAAAAA 955

RESULT 9
US-09-296-715-25
Sequence 25, Application US/09296715
Patent No. 6171839
GENERAL INFORMATION:
APPLICANT: MCGONIGLE, BRIAN
APPLICANT: O'KEEFE, DANIEL
TITLE OF INVENTION: SOYBEAN GLUTATHIONE-S-TRANSFERASE
TITLE OF INVENTION: ENZYMES
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: E.I. DU PONT DE NEMOURS AND COMPANY
STREET: 1007 MARKET STREET
CITY: WILMINGTON
STATE: DELAWARE
COUNTRY: UNITED STATES OF AMERICA
ZIP: 19898
COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE, 3.50 INCH
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: MICROSOFT WORD FOR WINDOWS 95
SOFTWARE: MICROSOFT WORD VERSION 7.0A
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/296,715
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: FLOYD, LINDA AXAMETHY
REGISTRATION NUMBER: 33,692
REFERENCE/DOCKET NUMBER: CL-1108
TELECOMMUNICATION INFORMATION:
TELEPHONE: 302-892-8112
TELEFAX: 302-773-0164
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 991 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
TISSUE TYPE: SOYBEAN
IMMEDIATE SOURCE:
CLONE: SSM.PK0067.G5
US-09-296-715-25

Query Match 3.7%; Score 42.4; DB 3; Length 991;
Best Local Similarity 54.5%; Pred. No. 0.19;
Matches 85; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

QY 980 AGTTGAGGTTATCATGTAACATAAGTTTCTTTAGGGCGCTGTTTCTATTGTCAA 1039
DB 800 AGTATGATTTTGTGGGAAACAATTATCTTGTGTGAGCAAGATGTTCTGTTTAA 859
1040 TATCATCAGCTTTGTGCTTATGCTTTTACAACTTATATGTACAACCTTTAAGTCA 1099
DB 860 TTTAATGACGCTGATTTGGTATGGCTATTTTAATTTTAACATAAAAAAGTGT 919
QY 1100 CCTCTTGGCTTACAAAAA 1135
DB 920 CAGTTTAAAAA 955

RESULT 10
US-09-832-129-13
; Sequence 13, Application US/09832129
; Patent No. 6936691
; GENERAL INFORMATION:
; APPLICANT: Fiascella et al.
; TITLE OF INVENTION: 19 Human secreted proteins
; FILE REFERENCE: P2045p1
; CURRENT APPLICATION NUMBER: US/09/832,129
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: PCT/US00/28664
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/163,085
; PRIOR FILING DATE: 1999-11-02
; PRIOR APPLICATION NUMBER: 60/172,411
; PRIOR FILING DATE: 1999-12-17
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 3265
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-832-129-13

Query Match 3.7%; Score 42; DB 3; Length 3265;
Best Local Similarity 59.0%; Pred. No. 0.42;
Matches 72; Conservative 0; Mismatches 50; Indels 0; Gaps 0;
QY 1014 TTAGGCGCTGGTTTCTTATGTCATATCATGCTTTGTGCTTATGTTTACAA 1073
DB 3131 TTAAACCTTCTTTCTTTCTTTATTAATGACCGCTTTTGCTTTCATTGTTACAG 3190
QY 1074 CTATATGTACAACTCTTAAAGTCACCTCTTGTCTTACAAAAA 1133
DB 3191 TTCTATTTTGAATAAATGTTCTCTCTTAAAAAAA 3250
QY 1134 AA 1135
DB 3251 AA 3252

RESULT 11
US-09-790-988-1/c
; Sequence 1, Application US/09790988
; Patent No. 6632935
; GENERAL INFORMATION:
; APPLICANT: SHIGENOBU, SHUJI
; APPLICANT: MATANABE, HIDEMI
; APPLICANT: HATTORI, MASAHIRA
; APPLICANT: SAKAKI, YOSHIYUKI
; TITLE OF INVENTION: GENOME DNA OF BACTERIAL SYMBIONT OF APHIDS
; FILE REFERENCE: 081356/0159
; CURRENT APPLICATION NUMBER: US/09/790,988
; CURRENT FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: JP2000-107160
; PRIOR FILING DATE: 2000-04-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 640681
; TYPE: DNA
; ORGANISM: Buchnera sp.
US-09-790-988-1

Query Match 3.7%; Score 41.6; DB 3; Length 640681;
Best Local Similarity 57.8%; Pred. No. 5.5;
Matches 74; Conservative 0; Mismatches 54; Indels 0; Gaps 0;
QY 968 TTGAATTCAGAGCTTGAGGTATCATGTAAACATAGTTTCTTAGGCGCTGGTT 1027
DB 527360 TAGATAAAATGTTTGTGTTTGTGATCTAAAAATATTTTGTGATACATTGATTT 527301

QY 1028 TTCTATGTCAATATCATGCTTTGTGCTTATGTTTACAACTTATATTCACA 1087
DB 527300 ATTCAATAGTCAAAATGATTAATCATTTATGTTATATAAAAAATTTAAAGATTATATA 527241
QY 1088 CTCCTTAA 1095
DB 527240 TTGATTA 527233

RESULT 12
US-10-033-301-23
; Sequence 23, Application US/10033301
; Patent No. 6930172
; GENERAL INFORMATION:
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Collin K.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2930R1c6
; CURRENT APPLICATION NUMBER: US/10/033,301
; CURRENT FILING DATE: 2001-12-27
; PRIOR APPLICATION NUMBER: 60/095,325
; PRIOR FILING DATE: 1998-08-04
; PRIOR APPLICATION NUMBER: 60/112,851
; PRIOR FILING DATE: 1998-12-16
; PRIOR APPLICATION NUMBER: 60/113,145
; PRIOR FILING DATE: 1998-12-16
; PRIOR APPLICATION NUMBER: 60/113,511
; PRIOR FILING DATE: 1998-12-22
; PRIOR APPLICATION NUMBER: 60/115,558
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/115,565
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/115,733
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/119,341
; PRIOR FILING DATE: 1999-02-09
; PRIOR APPLICATION NUMBER: 60/119,537
; PRIOR FILING DATE: 1999-02-10
; PRIOR APPLICATION NUMBER: 60/119,965
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: 60/162,506
; PRIOR FILING DATE: 1999-10-29
; PRIOR APPLICATION NUMBER: 60/170,262
; PRIOR FILING DATE: 1999-12-09
; PRIOR APPLICATION NUMBER: 60/187,202
; PRIOR FILING DATE: 1999-06-02
; PRIOR APPLICATION NUMBER: PCT/US99/12252
; PRIOR FILING DATE: 1999-06-02
; PRIOR APPLICATION NUMBER: PCT/US99/28634
; PRIOR FILING DATE: 1999-12-01
; PRIOR APPLICATION NUMBER: PCT/US99/28551
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US00/03565
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: PCT/US00/05841
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: PCT/US00/08439
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: PCT/US00/14941

PRIOR FILING DATE: 2000-05-30
PRIOR APPLICATION NUMBER: PCT/US00/15264
PRIOR FILING DATE: 2000-06-02
PRIOR APPLICATION NUMBER: PCT/US00/32678
PRIOR FILING DATE: 2000-12-01
NUMBER OF SEQ ID NOS: 38
SEQ ID NO 23
LENGTH: 2186
TYPE: DNA
ORGANISM: Homo sapiens
US-10-033-301-23

Query Match 3.6%; Score 40.4; DB 3; Length 2186;
Best Local Similarity 53.9%; Pred. No. 0.96; Mismatches 71; Indels 0; Gaps 0;
Matches 83; Conservative 0;

QY 982 TTGAAGTTATCATTAACATTAAGTTTCTTGAAGGCTGTTTCTATATGCAATA 1041
DB 2001 TCTGGCGGACACATGAAATGATATACCTTTATGCTGTGTTTTATCTTGAT 2060
QY 1042 TCATCAGCTTTGTGCTTATGCTTTTACAACTTATATGACAACTTAAAGTCAC 1101
DB 2061 ACATTGATTTTTCACGTAAGTCCACATATCTTCTATAGAGCGGTGATATATA 2120
QY 1102 TCTTGCTTACAAAAA 1135
DB 2121 GGGTTATGAGAAAAA 2154

RESULT 13

US-09-008-697A-11
Sequence 11, Application US/09008697A

PATENT NO. 6197504
GENERAL INFORMATION:
APPLICANT: Chow, King Lau
TITLE OF INVENTION: USES OF MAB-21
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham LLP
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/008,697A
FILING DATE: January 19, 1998
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Chan, Albert Wei-Kit
REGISTRATION NUMBER: 36,479
REFERENCE/DOCKET NUMBER: 50752-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0525
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 1278 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-09-008-697A-11

Query Match 3.5%; Score 40; DB 3; Length 1278;
Best Local Similarity 59.8%; Pred. No. 0.97;
Matches 67; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

QY 1024 GTTTCATATGTCATATCATCAGCTTTGTCCTATAGGTTTACAACTTATATGT 1083
DB 1166 GTTGTATTTTATTAATATCAATTCCTCAACCTTTTCACATCACTGATACGT 1225
QY 1084 ACAACTCTTAAAGTCACCTCTTGGCTTACAAAAA 1135
DB 1226 TCATCTTATTAATAATTAATTTTCTCCAAAAA 1277

RESULT 14

US-08-148-209A-1
Sequence 1, Application US/08148209A

PATENT NO. 5556780
GENERAL INFORMATION:
APPLICANT: Dzaou, Victor J
APPLICANT: Mukoyama, Masashi
TITLE OF INVENTION: TYPE-2 ANGIOTENSIN II RECEPTOR AND GENE
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: FLEHR, HOBBACH, TEST, ALBRITTON & HERBERT
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-4187

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/148,209A
FILING DATE: 05-NOV-1993
CLASSIFICATION: 800

ATTORNEY/AGENT INFORMATION:
NAME: Rowland, Bertram I
REGISTRATION NUMBER: 20,015
REFERENCE/DOCKET NUMBER: A-58491-1/BIR
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
TELEX: 910 277299

INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2862 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 132..1223

US-08-148-209A-1

Query Match 3.5%; Score 40; DB 2; Length 2862;
Best Local Similarity 61.5%; Pred. No. 1.4; Mismatches 40; Indels 0; Gaps 0;
Matches 64; Conservative 0;

QY 1032 ATTGTCATATCATCAGCTTTGTCCTATAGTTTTACAACTTATATGTAACACTCT 1091
DB 2747 ATAGCTAAATATATGATATGATGCTTTGTGTTTACGAATTTTACAAATATCG 2806
QY 1092 TTAAGTCACCTTTGCTTACAAAAA 1135
DB 2807 TAAATTAATCATTAATGCGGAAAAA 2850

RESULT 15

US-09-881-556A-1
Sequence 1, Application US/09881556A

PATENT NO. 6808900
GENERAL INFORMATION:
APPLICANT: Simonsen, J. Neil

```

: TITLE OF INVENTION: Cryptosporidium Parvum Antigens, Antibodies Thereto and Diagnostic
:
: TITLE OF INVENTION: Therapeutic Compositions Thereof
:
: FILE REFERENCE: SASK-008/01US
:
: CURRENT APPLICATION NUMBER: US/09/881,556A
:
: CURRENT FILING DATE: 2001-06-14
:
: PRIOR APPLICATION NUMBER: US 60/212,083
:
: PRIOR FILING DATE: 2000-06-15
:
: NUMBER OF SEQ ID NOS: 6
:
: SOFTWARE: PatentIn version 3.1
:
: SEQ ID NO 1
:
: LENGTH: 1380
:
: TYPE: DNA
:
: ORGANISM: Cryptosporidium parvum
:
: US-09-881-556A-1

```

Query Match	3.5%	Score 39.6	DB 3	Length 1380
Best Local Similarity	57.1%	Pred. No. 1.3		
Matches	72	Conservative	0	Mismatches 54
				Indels 0
				Gaps 0

Qy	1010	TTCTTTAGGGGCTGGTTTTCTCATGTGCAATATCATCAGCTTTGGTCTATGGTTTTA	10659
Db	1239	TTAATTAAGGGGTAGATTAAGATATTTTTCAATGTAATAAATTAAGCTTATTAAGAGATTAT	12988
Qy	1070	CAACTTATATTTCACACTCTTTAAGCACCCTCTTGCTACAAAAA	11229
Db	1289	TCGAATTAATTAATTAAGAGATATTAAGTTAGTTTATTTAAAAA	13588
Qy	1130	AAAAAA	1135
Db	1359	AAAAAA	1364

Search completed: March 20, 2006, 14:05:56
Job time : 229.218 secs

MADE THE BOARD

THIS PAGE LEFT BLANK